

The design and analysis of experiments with a second phase in the laboratory

C. J. Brien¹, B.D. Harch² and R.L. Correll³

¹School of Mathematics and Statistics, University of South Australia,
North Terrace, Adelaide, SA 5000, Australia.

²CSIRO Mathematical and Information Sciences,
120 Meiers Rd, Indooroopilly, Qld 4068, Australia.

³Rho Environmetrics,
PO Box 366, Highgate, SA 2063 Australia.

(E-mail: *chris.brien@unisa.edu.au*).

The paper aims to provide a systematic approach to designing the laboratory phase, or second part, of two-phase experiments, by outlining general principles. Two-phase experiments occur widely, although their two-phase nature is often not recognized. The need to randomize the material produced from the first phase in the laboratory phase is emphasized. Bases for categorizing designs are suggested and the use of decomposition tables to evaluate their properties discussed. Their analysis using analysis of variance and mixed-model estimation and taking account of the confounding is also considered. The methods are illustrated using examples.

Key Words: Analysis of variance; Laboratory experiments; Mixed models; Multiphase experiments; Multitiered experiments; Two-phase experiments.

1. Introduction

It is common for the material produced during an experiment to be processed in a laboratory. Reasons for this include the need to measure chemical attributes using equipment such as spectrometers, gas chromatographs or pH meters, to test using laboratory equipment, or to produce processed products such as wine, bread and malt that are subsequently assessed often by an expert panel. Such experiments will consist of two phases, in the sense of McIntyre (1955), with an experimental design required for each phase. Agricultural experiments, with a field and a laboratory phase, are a common example of two-phase experiments. Clinical trials can also result in two phases, clinical and laboratory, when specimens from patients are processed in a laboratory. For some experiments both phases occur in the laboratory, such as in food processing when there is a phase in which mixtures are prepared, and a processing phase to produce the final product. More generally experiments may be multiphase. We confine ourselves to those two-phase experiments that involve a first phase followed by a laboratory phase. Here, the laboratory assessment is to be interpreted broadly as a phase in which further processing, measurement, testing and so on are performed, even if, strictly speaking, a laboratory is not involved.

Two-phase experiments were first described by McIntyre (1955), although he considered only those whose analysis could be performed on means from the second phase following the first-phase design. Curnow (1959) showed that such an analysis was not appropriate for McIntyre's complicated virus example. The crucial feature that McIntyre incorporated into his designs was the use of a randomized design in each phase. Cox (1958, p.83) also pointed out that 'It is frequently not good enough to randomize just one stage [phase] of the experimental procedure and to leave the treatments systematically arrayed at other stages [phases]'. Two-phase experiments involve multiple randomizations (Brien and Bailey, 2006) and belong to the class of multitiered experiments (Brien, 1983). Wood et al. (1988) discussed their analysis, while a general approach to the analysis of

multitiered experiments has been provided by Brien and Payne (1999). They, along with Brien, May, and Mayo (1987), used examples of two-phase sensory evaluation experiments. Brien, Harch, and Correll (1998) presented a discussion of the design and analysis of the experiment described in Section 2. Smith et al. (2001) discussed the design and analysis, using mixed-model estimation, of wheat experiments that involved a field and milling phase, and in which not all units from the field phase are processed in the milling phase. Cullis et al. (2003) described the design and analysis, using mixed-model estimation, of a three-phase experiment that involved a field phase in which barley lines were grown, a malting phase in which barley malts were produced and a measurement phase in which several traits were determined. Again not all units from one phase continue on to subsequent phases. Kerr (2003) noted that microarray experiments can be the measurement phase of a two-phase experiment. Brien and Bailey (2006, Examples 1, 4, 9, 12, 14 and 15, and Figure 7) gave examples of experiments involving a first phase followed by a laboratory phase.

While it is clear that two-phase experiments with a laboratory phase occur in a wide range of situations, there has not been a comprehensive treatment of the general principles for designing experiments of this class. The purpose of this paper is to distill and emphasize those design principles and to demonstrate their application by example. In Section 2 an example is described, while in Section 3 general principles for designing such experiments are outlined. In Section 4 we give examples of two-phase experiments that employ equireplicate designs in both phases. In Section 5 we summarize the major issues that arise from the examples.

2. A biodiversity experiment

Harch et al. (1997) described a biodiversity experiment that started with a field experiment to look at the effect of two tillage treatments on bacterial and fungal diversity. It employed a randomized complete block design with four blocks. For each plot, two soil samples were taken at two different depths (0–5 cm versus 5–10 cm) at the one location in the plot. The resulting 32 soil samples were analyzed in a systematic order in 32 runs of a gas chromatograph and then reanalyzed again in that same order. In this example we refer to the laboratory unit as a run. Thus, there are 64 runs grouped into two Occasions (or two Int1s as introduced Section 4.1) at each of which the 32 soil samples were assayed. Now the two samples taken at a single depth were preprocessed using two different methods (ground versus sieved). Each Occasion was divided into two Intervals during each of which the 16 samples from two of the four blocks were assayed. The order of processing of the 16 samples from two Blocks, A and B say, is shown in the columns headed Method, Block, Plot and Depth in Table 1. Also given, in observation order, are the values for the response variable, a Gini coefficient computed from readings from BIOLOG™ plates at selected incubation times during a gas chromatograph run.

PLEASE INSERT TABLE 1 NEAR HERE

This arrangement has the obvious defect that Method and, to a lesser extent, Depth would be confounded with systematic trends across the Runs due to problems such as equipment drift, operator learning and fatigue and changes in the laboratory ambience. Furthermore, the seemingly most important and possibly smallest effect of Tillage is not confounded with potentially the smallest source of variation, that between pairs of consecutive runs.

3. General Principles

The notation used in this and subsequent sections is that $A \wedge B$ denotes a generalized factor with ab levels formed from the combinations of the a levels of A with the b levels B, $A \# B$ the interaction of A and B and $C[A \wedge B]$ the differences between C nested within the combinations of A and B.

3.1 DESIGNING THE LABORATORY PHASE

3.1.1 Some considerations

The first and laboratory phases of an experiment are often dealt with entirely separately. The use of statistical design principles in the first phase is well-appreciated, but the need to employ them in the laboratory is usually overlooked, in spite of the early work of McIntyre (1955) and Cox (1958). It has been common practice to process in a systematic order, for example field produce is processed in ‘field order’ in the laboratory, or processing order is not considered at all. It cannot be emphasized enough that both phases need to be planned together and that, in the laboratory phase, first-phase units should be randomized to second-phase units to avoid systematic trends affecting effects of interest. Hence such two-phase experiments should necessarily involve two-randomizations. As a result, they will involve three sets of factors or tiers (Brien, 1983; Brien and Bailey, 2006, Section 2) that, in general, can be thought of as the sets of factors indexing i) first-phase treatments, ii) first-phase units, and iii) laboratory units.

The laboratory-phase units are often the time at which an analysis is performed, or a position in a machine at a particular time when a set of specimens are processed together. Because of this we will refer to these units as positions which may be over time and/or space. Then blocking, to minimize the random variation affecting treatments, will involve grouping together positions in time (and space) to form blocks of homogeneous positions. The most obvious way to do this over time is to form laboratory blocks from consecutive times or time periods, but what structure should be placed on the positions? The innate structure will often be that positions and blocks are crossed so that a nonhierarchical design based on rows and columns could be employed. This would then result in the elimination of both smooth and nonsmooth trends across the positions, provided these trends are reasonably consistent between blocks. However, if consistent differences between positions across blocks are not expected, then a hierarchical structure is appropriate. Thus, what structure is adopted in a particular instance will in part depend on the innate structure and on the variation anticipated in the experiment. Example 9 of Brien and Bailey (2006) employs a nonhierarchical design because positional differences are to be isolated.

The next task that faces the designer is how to randomize the first-phase factors onto the structure for the laboratory phase. Designing the laboratory phase is, in a general sense, the same as designing single-randomization experiments: one needs to replicate, randomize and block positions and standard designs are often employed to do this. Similarly, the objective will be to confound first-phase, treatment effects with the smallest source of laboratory variation possible. It may be necessary to use the split-plot principle to confound different effects with different laboratory sources. The restrictions eventually decided upon for this randomization will be reflected in the crossing and nesting relationships between factors in the laboratory phase; these relationships will be exhibited in the laboratory panel of a randomization diagram for the experiment (Brien and Bailey, 2006). However, the number of factors to be randomized in the laboratory phase is generally larger than in a single-randomization experiment. Consequently, a more complicated design is likely to be needed and pseudofactors will more often be required to achieve a design that is balanced (Brien and Bailey, 2006, Sec. 8.2). Even more important is that it is impossible to observe all combinations of the levels of the factors being randomized in the second randomization. It is not surprising then, that there is a tendency to ignore some first-phase, unrandomized factors when it comes to randomizing the laboratory phase. For example, to simplify the laboratory randomization for a randomized complete block design, Plots is often ignored and just the randomization of Blocks and Treatments considered. The difficulty with this is that one can easily lose track of how sources of variation in the first phase affect the data that is eventually obtained, especially with more complicated designs. To avoid this, the approach to laboratory-phase randomization should *always* be to randomize *all* the first-phase, unrandomized factors (Blocks and Plots) to the laboratory units (positions), with the

factors randomized to the first-phase units (Treatments) taken into account where necessary.

A common practice is to combine material from the first phase for processing in the laboratory phase. However, as asserted by McIntyre (1955), it is important that product from each first-phase unit be separately evaluated in the second phase. Otherwise, sources of variation from the first phase may be inestimable and may lead to variance underestimation or the lack of an appropriate variance estimate.

3.1.2 Categories of designs

The following four features are proposed for characterizing differences between prospective designs:

1. Relationships between unrandomized factors in the laboratory phase: *hierarchical* or *nonhierarchical*, as mentioned above.
2. Type of two-phase randomizations: *two composed*, *two randomized-inclusive* or *multiple first-phase randomizations* (Brien and Bailey, 2006), where *two-phase randomizations* are comprised of the first-phase randomization(s) that randomize factors to first-phase units and the laboratory-phase randomization that randomizes first-phase units to laboratory units.
3. The introduction of *additional treatments at the laboratory phase*, or not.
4. The inclusion of *laboratory replicates*, or not.

In order to identify situations in which composed randomizations are feasible, different types of matching between factors from different tiers are defined. Matching applies to the number of levels combinations of a *set* of one or more factors that are to be randomized together to the same unrandomized factor(s). *Primary matching* occurs when this number divides the number of levels of each of the unrandomized factors to which the set is randomized. *Nested matching* occurs when the number of levels of the generalized factor formed from the set, and all factors nesting at least one of them, divides those for the generalized factor(s) formed from the unrandomized factors as follows:

- (a) when the set is completely randomized to the levels combinations of one or more unrandomized factors: the generalized factor formed from these unrandomized factors, and all factors that nest at least one of them;
- (b) otherwise, a generalized factor is formed for each unrandomized factor to which the set is randomized, provided it does not nest other such unrandomized factors; each generalized factor is formed from the unrandomized factor and all factors nesting it.

Otherwise, the set is *not matched* with the factors to which it is to be randomized.

Two composed randomizations are the simplest for two-phase experiments. To see when they apply, consider the sources derived from first-phase units that have treatment sources confounded with them. If, for each of these sources, the whole source is confounded with the same laboratory source(s), then the randomizations will be composed. It is conjectured that this will be feasible when there is nested matching between a) those unrandomized factors from the initial phase, that have had factors randomized to them, and b) the unrandomized factors from the laboratory phase. Orthogonal designs are feasible when there is also primary matching. Otherwise, nonorthogonal designs will be required and, in some cases, composed randomizations may not be possible due to the unavailability of a suitable design. The considerable advantage of composed randomizations is that the randomization of the initial phase can be ignored in performing the second randomization. An even more substantial advantage is that the properties of the design used in the initial phase are not affected by the design used for the laboratory phase, as discussed below.

Figure 1 shows the randomization diagram for the composed randomizations of a two-phase experiment in which a randomized complete block design was used in the first phase and a balanced incomplete block design in the second phase. For the first randomization, there is primary and nested matching of Treatments and Plots. For the second randomization, as it is a balanced incomplete block design, $st > k$ and $rst = kb$ where r is the number of replicates of the plots in the second phase. So there must be nested matching of Blocks^Plots with Runs^Positions as st obviously divides kb ; we do not need to check the nested matching with Runs, because it nests Positions. There cannot be primary matching of Blocks^Plots with Positions. Hence, the composed randomizations were feasible, but not using an orthogonal design.

PLEASE INSERT FIGURE 1 NEAR HERE

Consider Example 4 of Brien and Bailey (2006) in which treatments are randomized to fibres and fibres are randomized to tests. In all three plans Plots, that has Treatments randomized to it, is randomized to Tests and there is both primary and nested matching of Plots to Tests. All three plans involve composed randomizations, but, only in Plan A is the whole of each of the three field sources confounded with the same laboratory source: the whole of Blocks is confounded with Operatives and the whole of both Plots [Blocks] and Fibres [Blocks \wedge Plots] are confounded with Tests [Operatives]. In Plans B and C the source Fibres [Blocks \wedge Plots] is split; part is confounded with Operatives and the rest with Tests [Operatives]. The need for this can be predicted from the observation that the matching of Fibres to Operatives is not nested. However, the two plans still involve composed randomizations because the whole of the source Plots [Blocks], with which Treatments is confounded, is confounded with the one source Tests [Operatives]. What happens with Fibres is irrelevant as it has nothing confounded with it in the first phase.

The simplest approach outlined by McIntyre (1955) for designing the laboratory phase is to use the same design in both phases so that each unrandomized factor from the first phase has an equivalent factor in the laboratory phase. Clearly, there will be both primary and nested matching of the unrandomized factors from the first phase to those from the laboratory phase, and the randomizations will be composed.

Two randomized-inclusive randomizations have the same objective as composed randomizations: the randomization of treatments to first-phase units and of these units to laboratory units. The difference is that randomized-inclusive randomizations are used when, in randomizing first-phase units to laboratory units, it is necessary to split up sources derived from first-phase units that have sources confounded with them. Such sources would be kept whole in composed randomizations. It is conjectured that randomized-inclusive randomizations will be required when the nested factor matching required for composed randomizations does not pertain. This lack of matching is signalling the need to split one or more sources that were unrandomized in the first-phase. Pseudofactors are used to do this by creating subgroups of a factor that split the requisite source(s). However, to ensure that the way in which the split is done results in good properties for the randomized sources from the first phase, the randomized factors from the first phase must be taken into account. A serious disadvantage of randomized-inclusive randomizations is that the design chosen for the laboratory design may degrade the properties of the design for the initial phase.

One feature of experiments involving a laboratory phase that may make it necessary to use randomized-inclusive, rather than composed, randomizations is that specimens must be processed in groups whose size mandates their use. For example, the laboratory equipment performs a fixed number of analyses jointly in a run and it is desirable that each run forms a block. Even so, two composed randomizations will be feasible if there is nested matching and a balanced design is available for the second randomization as in Figure 1. The Beetle damage example in Section 4.2 is one with constraints on group size in the laboratory phase that results in randomized-inclusive randomizations. It also demonstrates that in some cases of randomized-inclusive randomization, the design

for the first phase must fit with the requirements of the laboratory phase if a good design is to be obtained for the laboratory phase. Indeed, in the example, the laboratory-phase design was obtained first and the first phase derived from it. Consequently, it is dangerous to ignore the laboratory phase until after the first phase.

Multiple first-phase randomizations are uncommon in two-phase experiments. An example would be where the initial phase involves a superimposed experiment that employs randomized-inclusive randomizations. A third randomization is needed for the laboratory phase and this may be in the nature of either a composed or randomized-inclusive randomization. It depends on whether unrandomized sources from the first phase, that had factors randomized to them, are to be split.

When *treatments are added in the laboratory phase*, then additional randomizations are required and independent or coincident randomizations (Brien and Bailey, 2006) are likely to be involved. Example 12 from Brien and Bailey (2006) provides an example of this. The two-phase randomizations of harvesters to lots and lots to plates are composed. The addition of treatments in the laboratory phase adds a randomization, that of treatments to plates. It is coincident with part of the randomization of lots to plates.

It is common to include *laboratory replicates*, although these are unnecessary if the variability in the first phase is substantially greater than that in the laboratory phase.

3.1.3 The properties of designs

To assess a proposed design, the randomization-based decomposition table (Brien and Bailey, 2007) will be used to provide a summary of its properties: the confounding, efficiency factors and residual degrees of freedom. We strongly recommend that this table is always obtained as a check on the properties of the proposed design, even when the data is not going to be analyzed by analysis of variance. This is particularly easy to do for structure-balanced experiments using the AMTIER procedure, distributed with Edition 9 of GenStat (Brien and Payne, 2006); it will produce such a table from just the factors, with no need to specify a response variable.

In determining the properties of the design for a multiphase experiment, there commonly arises the question as to whether one of the phases determines the form of the analysis. McIntyre (1955) asserted for his designs that the analysis is determined by the first phase. This is indeed true for the special case in which the numbers of first-phase and laboratory units are the same, irrespective of whether composed or randomized-inclusive randomizations are employed. Then Lemma 4.2 of Brien and Bailey (2007) applies so that the decomposition obtained by combining that for the first-phase and laboratory units is the same as that for the first-phase units. This must occur when there are no laboratory replicates, as in the example in Section 4.2; it will also occur when there are laboratory replicates, provided the laboratory replicates involve the processing of different portions from the produce of a first-phase unit and a factor is included to account for these portions (see Brien and Bailey, 2006, Example 12). However, in general, both phases may play a role, although that of the laboratory phase may be quite limited. Certainly, the design used in the initial phase will have certain properties. However, the laboratory phase will not affect these properties if composed randomizations are involved. On the other hand, if randomized-inclusive randomizations are used, it is possible that the design for the laboratory phase may diminish the properties of the design for the initial phase; it cannot improve them.

The laboratory-phase design has no effect for composed randomizations because i) the whole of each source from the initial phase is estimated from the same sources from the laboratory phase, and ii) as shown by Brien and Bailey (2007), the efficiency factors for the design for the initial phase multiply with those for the design for the laboratory phase. Hence, the greatest impact that the laboratory phase can have on the properties of the initial phase is that information may have to be combined from several sources. For the first-phase design in Figure 1, Treatments is orthogonal to

Blocks and the Residual degrees of freedom are $(s - 1)(t - 1)$. The laboratory-phase design does not alter this; it is just that first-phase sources will be confounded with both Positions and Runs. The three-tiered example, described by Brien and Payne (1999) and Brien and Bailey (2006, Figure 12), is a more complicated example.

In contrast, randomized-inclusive randomizations involve the splitting of some sources from the initial phase in randomizing them to the laboratory phase. Consequently, the first-phase design determines the maximum residual degrees of freedom for first-phase treatments. In the nonhierarchical design proposed for the Biodiversity example in Section 4.3 and the Beetle damage example in Section 4.2 the laboratory phase reduces the degrees of freedom of the Residual for sources from the first-phase.

3.2 ANALYSIS OF DATA FROM EXPERIMENTS

A fundamental problem in the analysis of experiments is to make sure that all the sources thought important enough to be allowed for in the randomization are included in the model for the data, without adding terms that are not. To ensure this, we advocate that the analysis be based on a full randomization-based mixed model: a mixed model that is obtained from the randomization model by modifying it as described by Brien and Bailey (2006, Section 7). To formulate this model, it is recommended that one starts with the randomization diagram (Brien and Bailey, 2006) for the experiment. These have the advantage that they show i) the experimental units by showing which factors are randomized to which and ii) the restrictions placed on the randomization in the form of the crossing and nesting relationships between the unrandomized factors. Then the diagram can be used in deriving the randomization-based decomposition table, that in turn is useful in formulating the model. In this we urge that the model not be distorted to produce merely a model of convenience, such as by the expedient substitution of factors from one tier for those in another. Even though this does reduce the number of factors that have to be specified, it can be misleading and means that the outcomes of the randomizations are neither retained nor incorporated into the analysis. In particular, a term that correctly identifies each experimental unit is not included in the analysis. For example, we would not substitute Treatments for Plots in the model for a randomized complete block design as it would replace $\text{Blocks} \wedge \text{Plots}$ with $\text{Blocks} \wedge \text{Treatments}$. Omitting this term means there is not a term identifying this experiment's only experimental unit and its replacement gives the incorrect impression that a block-treatment interaction is being allowed for in the model when the real intention is that a source of inherent variability be incorporated.

At least two methods can be used to analyse the data: analysis of variance and mixed-model estimation. The AMTIER procedure (Brien and Payne, 2006) is purpose-built to perform the analysis of variance of multitiered experiments that are structure balanced (Brien and Bailey, 2007); it uses the algorithm described by Brien and Payne (1999). Although it is possible to use any analysis of variance or regression software, in practice it is only feasible for orthogonal designs. Mixed-model estimation, based on Residual or Restricted Maximum Likelihood (REML) (Patterson and Thompson, 1971), is widely available including in GenStat (Payne et al., 2006) and SAS (Littel et al., 2006). It is particularly applicable for nonorthogonal experiments and those with pseudofactors, because combined estimates will be obtained and the specification of pseudofactors can usually be avoided. However, it has the disadvantage that, if there are totally aliased random terms, the model supplied to a procedure will not be the full randomization-based mixed model for the data, but a model of convenience that does not include some totally aliased terms. For example, when random terms are totally confounded by fixed effects they need to be removed for the fitting algorithms to work. In addition, when are laboratory random terms are totally confounded by first-phase random terms, then some must be omitted. The omission of random terms will certainly be necessary when, as discussed in Section 3.1.3, the numbers of first-phase and laboratory units are equal. Clearly, all

the laboratory-unit terms can be omitted. In addition, it is only possible to estimate the sum of the variability between individual first-phase units and that between individual laboratory units. Then, as most software does not require an explicit term for unit variability, usually both can be omitted. However, when randomized-inclusive randomizations are employed it may be desirable to include some laboratory-unit terms so that pseudoterms can be omitted. The example in Section 4.2 illustrates these points. Also, convergence problems may be experienced when some random terms are near zero or have low degrees of freedom. In addition, low degrees of freedom cause difficulties because of the asymptotic approximations on which the analysis is based.

A mixed model will be expressed in two factor formulae in which A/B indicates that B is nested within A, and A*B that A and B are crossed. The two formulae will be separated by a ':', that to the left specifying the fixed terms and that to the right the random terms. Each formula will be expanded using the rules of Wilkinson and Rogers (1973) yielding two sets of terms. In each set, the terms are linked by a '+' indicating an additive model and each term consists of a set of factor names separated by '^' that replaces the '.' of Wilkinson and Rogers (1973). Whether the order of the terms matters depends on the software and options used. The order of the fixed terms may be important where calculations are based on adjusting each term for those before them, while the order for random terms is generally unimportant. For multitiered experiments, we follow the practice of putting first-phase random terms before the laboratory random terms, because the former are more likely to correspond to proper subspaces of the second especially for equal numbers of first-phase and laboratory units.

4. Examples

Experiments with a hierarchical laboratory phase are discussed in Sections 4.1 and 4.2, while the example in Section 4.3 has a nonhierarchical laboratory phase. Other examples of two-phase experiments in which the laboratory phase is hierarchical are provided by Brien and Bailey (2006, Examples 1, 4, 12, 14 and Figure 7). Examples with a nonhierarchical laboratory phase are provided by McIntyre (1955, Example 4 and the main virus example) and Brien and Bailey (2006, Examples 9 and 15).

4.1 HIERARCHICAL LABORATORY PHASE FOR THE BIODIVERSITY EXPERIMENT

As a preliminary step in redesigning the experiment described in Section 2, the magnitudes of the sources of variation in the laboratory phase are investigated. Here we seek only to eliminate the sources of field variation and treatment differences, whose interpretation is obscured in any case by the lack of randomization. To do this, an analysis of variance, not randomization-based but utilizing a structure formula for each phase as in Brien and Payne (1999), was performed on the Gini coefficient multiplied by 100 using the **AMTIER** procedure in GenStat. For this analysis, the 64 runs are divided hierarchically into six sets of time intervals according to six two-level factors, named Int1 to Int6; the factor Int_{*j*} divides each interval for Int_{*(j-1)*} in half. The values of these factors are given in Table 1. Clearly Int3 corresponds to Methods as each Int3 interval consists of 8 consecutive runs during which 8 samples from the same Method were processed. Similarly, Int5 and Int6 correspond to Plots and Depths respectively. Int2 corresponds to the difference between Blocks 1 and 2 versus 3 and 4 while Int4 corresponds to the differences within pairs of Blocks.

The analysis-of-variance table is given in Table 2. To produce it requires the use of pseudofactors for Blocks and Samples that indicate which Blocks and Samples, respectively, were processed together. The pseudofactors will split the Block and Sample [Block \wedge Plot \wedge Depth] sources as shown in Table 2. The pseudofactor for Blocks, say B₁, will have two levels and observations will have the same level for it if they are from blocks in the same Int2 of an Int1. The pseudofactor for Samples,

say S_1 , will have two levels and observations will have the same level for it if they are from samples in the same Int3 of an Int1∧Int2 combination.

PLEASE INSERT TABLE 2 NEAR HERE

From the analysis, there is a suggestion that Int4 and Int5 are sources of extra variation in this experiment ($p < 0.25$). That is, pairs of consecutive runs are perhaps more homogeneous than groups of four or eight consecutive runs. It is also worth noting that the Residual mean squares for Int1, Int2 and Int3 are no larger than that for Int4 and so they would have zero or negative variance components. This indicates that a model with positive variance components is inappropriate for this data and that either a) the terms should not be included in the model or b) the components should not be constrained to be positive. One possible reason for these negative components is that the equipment is recalibrated at the beginning of each Int3 interval. If this was known to be the case, then a negative component for Int3 is appropriate, as is combining the three sources involving Int1, Int2 and Int3. The result will be a single negative component for the combined source Int1 ∧ Int2 ∧ Int3, which should not be combined with the Int4 source as this would result in the underestimation of the corresponding variance component with consequent effects as discussed by Littell et al. (2006, Section 4.7).

As suggested in Section 3.2, the analysis could also be obtained using mixed-model estimation. The following model, with terms based on the sources in Table 2 and with the components for i) B_1 and Block and ii) $B_1 \wedge \text{Sample}$ and $\text{Block} \wedge \text{Plot} \wedge \text{Depth} \wedge \text{Sample}$ constrained to be equal, will fit without aliased components:

$$\begin{aligned} \text{Tillage*Depth*Method} : B_1/\text{Sample} + ((\text{Block}/\text{Plot})*\text{Depth})/\text{Sample} - \text{Depth} \\ + \text{Int1}\wedge\text{Int2}\wedge\text{Int3}/\text{Int4}/\text{Int5}/\text{Int6}. \end{aligned}$$

Note the placement of random terms from the field phase before those involving Int factors, all of the former corresponding to proper subspaces of the latter as can be verified from Table 2. The advantages of mixed-model estimation, in this case, are that the pseudofactor for Samples is not required and combined estimates for each of Blocks and $\text{Block} \wedge \text{Plot} \wedge \text{Depth} \wedge \text{Sample}$ are obtained. However, when using GenStat to perform this analysis without the terms that involve B_1 , all the variance components based on the Int factors are reported as being aliased. This is in spite of there being Residual degrees of freedom for them all. It appears that the problem is the confounding of a field-phase source with more than one laboratory source; here, as is evident in Table 2, this has happened with both Block and Sample [$\text{Block} \wedge \text{Plot} \wedge \text{Depth}$] and causes a problem with estimating the Block effects.

The analysis in Table 2 begs the question as to what design, in hindsight, one might have used for this experiment. Essentially, we have the factors Tillage, Depths and Methods, yielding 8 treatments, that are of particular interest to the researchers and need to be randomized to the different runs sequences. Further, we take as unalterable that the equipment will be recalibrated after every 8 runs; a set of 8 consecutive runs will be called an Interval, and so Intervals corresponds to Int1 ∧ Int2 ∧ Int3. Also, the only designs to be considered are those where all 32 samples from the field phase will be processed once and then reprocessed a second time in a different order. Thus the basic set-up of the laboratory phase, to which field units are to be randomized, can be described as 2 Occasions × 4 Intervals × 8 Runs, as illustrated in Table 3a.

PLEASE INSERT TABLE 3 NEAR HERE

In the absence of any further information from the researchers, it would be ideal if laboratory blocks of 8 homogeneous runs could be identified to accommodate the 2 Plots × 2 Samples × 2 Depths in a field block. Our analysis in Table 2 indicates that groups of 8 Runs within an Interval

will not accomplish our aim. However, the four Intervals within an Occasion are similar, as are pairs of Runs, so that a group of two consecutive Runs across four Intervals would be homogeneous. The proposed framework would become that in Table 3b; the sequence of runs within an occasion would remain by rows. The randomization diagram for this design is in Figure 2. There is primary and nested matching of Blocks and Times and of the generalized factor $\text{Depths} \wedge \text{Plots} \wedge \text{Samples}$ with $\text{Intervals} \wedge \text{Runs}$. There is nested matching of the latter pair because the product of all samples factors (32) divides that of all runs factors (64). Hence, the two-phase randomizations of field treatments to samples and the latter to runs are composed. Besides this design involving a hierarchical laboratory phase and two-phase randomizations that are composed, there are treatments added in the laboratory phase and duplicates are included. The randomizations of field and lab treatments to samples are independent.

Table 4 gives the randomization-based decomposition table showing the confounding for the proposed randomization. Note that the Residual under Tillage, like that in Table 2, has three degrees of freedom. This is determined by the randomized complete block design used in the field phase.

Solely as a check on laboratory variation for the proposed design, the data is reanalyzed with the laboratory-phase structure for this design and the original treatment structure. This is reasonable here because Table 3b merely relabels the runs in Table 3a: the researcher would still do eight runs in each of four intervals on both occasions as in the actual experiment. Only the assignment of treatments within Occasions changes. The Residual mean square for $\text{Intervals} \wedge \text{Runs}$ [$O \wedge T$] of 1.09 is little different from the value for the last Residual in Table 2. The proposed design has the advantage that all the terms of interest are confounded with the laboratory source that is likely to have the smallest variability. Importantly, this means that it has maximum Residual degrees of freedom for testing terms involving Tillage, Methods and Depths which, although they are limited by the field-phase arrangement, have not been reduced by confounding field-phase terms with different laboratory-phase terms. Blocks are of no interest to the researcher and so their confounding with the more variable Intervals is of no consequence.

PLEASE INSERT FIGURE 2 AND TABLE 4 NEAR HERE

It is emphasized that the advantages of the arrangement that has been proposed here do rely on the similarities between different Intervals within an Occasion. If it happened that there was substantial variation between Intervals within an Occasion, then clearly groups of Runs within Intervals will be beneficial. One would need to choose between groups of size 2, 4 or 8, of which 2 and 4 seem preferable. Thus at least one effect will have to be randomized to a larger source of variation. Perhaps Depths is anticipated to have large differences and so is a candidate.

4.2 BEETLE DAMAGE EXPERIMENT

Peacock et al. (2003) described an experiment to investigate a field observation that the presence of beetle damage on willows might be inhibiting rust development. It involved a glasshouse and a laboratory phase and is used to provide the context for the current example, but with the details changed to provide an example with some novel features. Suppose that in the glasshouse phase there are 12 treatments, simulating different extents and timing of beetle damage, that are assigned to 60 locations each containing a plant. Because of limitations in the glasshouse, sets of 6 locations on a bench form blocks and a resolvable incomplete block design is to be used to assign treatments. To evaluate the damage treatments, suppose that a leaf disc is to be taken from the plant at each location so that there will be 60 discs to be processed in the laboratory phase.

PLEASE INSERT FIGURE 3 NEAR HERE

In the laboratory phase the discs are to be put onto 20 plates, the plates having just three cells so that the discs from the plants from only three Locations can be placed on each Plate. The plates were

divided into 5 groups and on each of 5 occasions all the discs from the same replicates were applied to the 4 plates in a group. The factor Locations is to be assigned to Plates and Cells. However, Locations and Plates are not matched because there are 60 combinations of Locations in Benches and Reps and this does not divide 20, the number of combinations of Plates in Occasions. Thus the source for Locations within Benches and Reps has to be split across some laboratory sources and pseudofactors will be needed for this. It is helpful that the number of locations per bench in the glasshouse phase (6) is twice the number in cells per plate (3) in the laboratory phase and that both of these divide the number of treatments (12). So, as well as the laboratory phase being hierarchical, the randomizations will need to be randomized-inclusive and there are no treatments introduced in the laboratory phase nor are there laboratory replicates.

The first step in producing a design for this experiment is to obtain a resolvable incomplete block design for $v = 12$, $k = 3$ and $r = 5$, which was done using CycDesigN (Whitaker, Williams, and John, 2002). The nonrandomized layout is given in Table 5. It is a nonorthogonal design for which the average efficiency (John and Williams, 1995) is 0.698, compared with the upper bound of 0.721. The Plates within an Occasion are intentionally ordered according to the lowest treatment for each Plate. This design will be used in both phases, but with different randomizations for each phase.

PLEASE INSERT TABLE 5 NEAR HERE

It will be used to assign the treatments to locations in the greenhouse phase by combining pairs of Plates, as shown in Figure 3, to form Blocks of size six. The pairing of plates was chosen so that four of the six pairs of the treatments 1, . . . , 4 occurred twice on the same Bench and the other two occurred once. This yields a resolvable incomplete block design for $v = 12$, $k = 6$ and $r = 5$ with average efficiency 0.893, compared with the upper bound of 0.898. The circle in Figure 3 signifies that a specific design, rather than complete randomization, was used in the assignment to Damages and the lack of a ‘ \perp ’ indicates that it was nonorthogonal.

In the laboratory phase the design in Table 5 will be used to assign Damages and Locations to Plates and Cells. Clearly, the numbers in Table 5 specify which Damages are assigned to the Cells within a particular Plate on one Occasion. But what about Locations? To ensure that it is included in the analysis we need to realize that the numbers in Table 5 can be interpreted as an indirect reference to the Locations with the specified Damages. Then, two pseudofactors, L_1 with two levels and L_2 with three levels, can be constructed and these amount to a relabelling of Locations. The first has the same level and the second different levels for the three Locations that are to be on the same Plate in an Occasion. They can be used to randomize the locations to cells as depicted in Figure 3; the ‘ \blacklozenge ’ in this figure indicates that Damages determines the pseudofactors. This second view of the randomization focuses on randomizing Locations in such a way that Damages is also suitably randomized. Clearly, in order to form the pseudofactors for Locations, the result of the randomization of treatments to locations must be known. Hence, the randomizations are randomized-inclusive.

To further examine the properties of the proposed design its decomposition table is given in Table 6. Clearly the structure on locations, extended to include the pseudofactors, is orthogonal in relation to that on cells, a reflection of the close relationship between the blocking in the glasshouse and laboratory phases. However, unlike the examples in Brien and Bailey (2006), the structure on treatments is not structure balanced in relation to the combined decomposition of that on cells and locations. Consequently, the preferred method of analysis is mixed-model estimation. Another point about this experiment is that, if there were a response from the first phase, the degrees of freedom for the Residual for Locations [Reps \wedge Benches] would be 39; this has been reduced effectively to 29 degrees of freedom in the laboratory phase, although it cannot be separated from variability arising from Cells [Occasions \wedge Plates]. Also, this example shows that nonorthogonal designs, at least for treatments in the first phase, can be used; this has not been previously demonstrated in the literature.

PLEASE INSERT TABLE 6 NEAR HERE

A full randomization-based mixed model, with locations sources before cells source, for analyzing this experiment is: ‘Damages : Repls/Benches/Locations + Occasions/Plates/Cells’. However, it is clear that, due to the confounding in the experiment, there are two pairs of variance components for which only the sum of the pair’s components can be estimated: $\sigma_{\text{OPC}}^2 + \sigma_{\text{RBP}}^2$ and $\sigma_{\text{O}}^2 + \sigma_{\text{R}}^2$ (see Table 6). As noted in Section 3.2, the first pair can be foreseen because the numbers of locations and cells are equal (60). One model of convenience is: ‘Damages : Repls/Benches + Occasions^Plates + Repls^Benches^Locations’. It does not include all sources of variation in the experiment but will provide estimates of the estimable quantities. Other models can be used, but they must include Repls^Benches and Occasions^Plates.

4.3 NONHIERARCHICAL LABORATORY PHASE FOR THE BIODIVERSITY EXPERIMENT

As discussed in Section 4.1, Table 3a gives the basic set up for the laboratory phase for the biodiversity experiment described in Section 2. It consists of two Occasions in each of which there are 4 Intervals of 8 Runs. Because Intervals and Runs are inherently crossed, it is natural to consider an orthogonal design with crossed rows and columns to assign the 4 Blocks \times 2 Plots \times 2 Depths \times 2 Samples to the runs, as indicated in Figure 4. Runs in this design will be latinized (Williams, 1986) across Occasions, because it is crossed with Occasions. However, there is neither primary nor nested matching of Depths^Plots^Samples with Intervals. As a result an experiment based on this design will differ from that in Section 4.1, not only in having a nonhierarchical laboratory phase, but also in requiring randomized-inclusive randomizations for the two-phase randomizations.

PLEASE INSERT FIGURE 4 NEAR HERE

The design key method (Patterson and Bailey, 1978) implemented in GenStat (Payne et al., 2006) was used to obtain a design. The selected design key, in terms of the first letters of the factors and subscripts indicating pseudofactors, was:

$$\begin{array}{ccccc} B_1 & B_2 & P_1 & D & S_1 \\ I_1 & I_2 & OI_1R_1 & OI_2R_2 & I_1I_2R_3 \end{array}$$

Pseudofactors, each with two levels, have been assigned to the non-prime factors Blocks, Intervals and Runs. Now we cannot ignore Tillage and Methods in this design, or their effects will be nonorthogonal to terms involving the other factors. The factors Plots and Samples have also been replaced by pseudofactors, that amount to a relabelling of the factors. The pseudofactor P_1 identifies Plots with the same Tillage and so is the same as Tillage; similarly with S_1 for Samples and Methods. The pseudofactors will be randomized in assigning samples to runs. Their use ensures a favourable randomization of the treatment factors is obtained. Of course, one could expediently replace P with T and S with M, but this is not done so as to retain Plots and Samples in the randomization and analysis. The aliases for P_1 and S_1 in the design key have been chosen so that they, along with all the interactions involving them and D except the three-factor interaction, are confounded with the Occasion^Interval^Runs term. Consequently, so are the terms involving Tillage, Method and Depth.

From the randomization-based decomposition table in Table 7 it is clear that the design is orthogonal, provided pseudofactors are used to identify samples sources that are confounded with more than one runs source. It is also concluded that, in order to use a nonhierarchical design for this experiment, we have to sacrifice i) one Residual degree of freedom for Plots [Blocks], ii) any valid information about Tillage # Methods # Depths, the latter because it is confounded with Runs, and iii) Residual degrees of freedom for Samples [Blocks \wedge Plots \wedge Depths].

PLEASE INSERT TABLE 7 NEAR HERE

As a check on whether the design proposed here might have lead to an improvement over the hierarchical design, the data were reanalyzed using the laboratory structure for the design and the original treatment structure. The final Residual mean square is 1.41 which is in excess of that for the proposed hierarchical design. It is concluded that the nonhierarchical design would not have done better in this experiment, but is included to demonstrate how such a design might be obtained.

5. Discussion

Anecdotally little attention is paid to the order in which material is processed in the laboratory. A major aim of this paper is to increase awareness of the need to randomize both the initial and laboratory phases and it is our hope that our systematic account of the design of such experiments will promote this. The examples show how the design is affected by whether the laboratory phase is hierarchical, whether composed or randomized-inclusive randomizations are used, whether treatments are added at the laboratory phase and whether laboratory replicates are included. Further, the example in Section 4.1 illustrates that the blocking that provides the smallest variability between individual units may not be the most obvious. On the other hand the example in Section 4.2 highlights that, while blocking in the first phase will often play a substantial role in the determination of the blocking in the laboratory phase, there are instances in which restrictions in the laboratory phase will need to be taken into account in the first phase.

The examples also demonstrate the usefulness of the decomposition table in formulating a mixed model for the experiment. In particular, in the Beetle damage example in Section 4.2, even though mixed-model estimation is its preferred method of analysis, the decomposition table gave insight into the correct parameterization of the model for the random effects by revealing the estimable quantities. It has been seen that often the full randomization-based mixed model will be modified either to remove terms totally aliased or confounded by other terms or, less acceptably, to substitute some terms for others so as to reduce the number of factors to be specified. The resultant model of convenience, while it may be useful for obtaining the analysis from a software package, must not be regarded as equivalent to the full randomization-based mixed model that includes *all* potential sources of variability for the experiment. Further, it is throwing away information about the actual randomization for the experiment. An example where it is virtually impossible to reduce the randomization-based mixed model is the three-tiered experiment given by Brien and Payne (1999); the only possible modification is that one of Occasions or Squares needs to be omitted. Also, the use of pseudofactors in randomized-inclusive randomizations is exhibited in the two examples that employed them. They are used to randomize unrandomized factors from the first phase to laboratory factors and, in doing this, they provide a connection between the randomized and unrandomized factors from the first phase. However, they are not needed for the analysis if mixed-model estimation is used.

Two issues that will be the subject of further work are designing for trend in the laboratory phase and the best way to disperse laboratory replicates in the laboratory phase.

ACKNOWLEDGMENTS

We are very grateful to the referees for suggestions that led to substantial reworking and improvement of an earlier version of the paper. We would also like to thank Clive Pankhurst and Bruce Hawke who provided us with the data for the Biodiversity example in Section 2 and to Gillian Arnold for bringing the Beetle damage example in Section 4.2 to our attention.

REFERENCES

Brien, C. J. (1983), "Analysis of variance tables based on experimental structure," *Biometrics*, 39, 51–59.

- Brien, C. J. and Bailey, R. A. (2006), “Multiple randomizations (with discussion),” *Journal of the Royal Statistical Society, Series B*, 68, 571–609.
- (2007), “Decomposition tables for multitiered experiments,” Submitted for publication.
- Brien, C. J., Harch, B. D., and Correll, R. L. (1998), “Design and ANOVA for experiments involving a field trial and laboratory analyses,” Paper presented to The Ninth International Conference on Quantitative Methods for the Environmental Sciences, Gold Coast, Australia.
- Brien, C. J., May, P., and Mayo, O. (1987), “Analysis of judge performance in wine-quality evaluations,” *Journal of Food Science*, 52, 1273–1279.
- Brien, C. J. and Payne, R. W. (1999), “Tiers, structure formulae and the analysis of complicated experiments,” *The Statistician*, 48, 41–52.
- (2006), “AMTIER Procedure,” in *GenStat Reference Manual Release 9, Part 3. Procedure Library PL 17*, eds. R. W. Payne and P. W. Lane, Hemel Hempstead: VSN International, pp. 73–75. URL: <http://www.genstat.co.uk>.
- Cox, D. R. (1958), *Planning of Experiments*, New York: Wiley.
- Cullis, B. R., Smith, A. B., Panozzo, J. F., and Lim, P. (2003), “Barley malting quality: are we selecting the best?” *Australian Journal of Agricultural Research*, 54, 1261–1275.
- Curnow, R. N. (1959), “The analysis of a two phase experiment,” *Biometrics*, 15, 60–73.
- Harch, B. D., Correll, R. L., Meech, W., Kirkby, C. A., and Pankhurst, C. E. (1997), “Using the Gini coefficient with BIOLOG substrate utilisation data to provide an alternative quantitative measure for comparing bacterial soil communities,” *Journal of Microbial Methods*, 30, 91–101.
- John, J. A. and Williams, E. R. (1995), *Cyclic and Computer Generated Designs*, 2nd ed., London: Chapman and Hall.
- Kerr, M. K. (2003), “Design considerations for efficient and effective microarray studies,” *Biometrics*, 59, 822–828.
- Littel, R., Milliken, G., Stroup, W., Wolfinger, R., and Schabenberger, O. (2006), *SAS for Mixed Models*, 2nd ed., Cary: SAS Press.
- McIntyre, G. A. (1955), “Design and analysis of two phase experiments,” *Biometrics*, 11, 324–334.
- Patterson, H. D. and Bailey, R. A. (1978), “Design keys for factorial experiments,” *Applied Statistics*, 27, 335–343.
- Patterson, H. D. and Thompson, R. (1971), “Recovery of inter-block information when block sizes are unequal,” *Biometrika*, 58, 545–554.
- Payne, R. W., Harding, S. A., Murray, D. A., Soutar, D. M., Baird, D. B., Welham, S. J., Kane, A. F., Gilmour, A. R., Thompson, R., Webster, R., and Tunnicliffe Wilson, G. (2006), *The Guide to GenStat Release 9, Part 2 Statistics*, Hemel Hempstead: VSN International. URL: <http://www.genstat.co.uk>.
- Peacock, L., Hunter, T., Yap, M. L., and Arnold, G. (2003), “Indirect interactions between rust (*Melampsora epitea*) and leaf beetle (*Phratora vulgatissima*) damage on *Salix*,” *Phytoparasitica*, 31, 226–235.

- Smith, A. B., Cullis, B. R., Appels, R., Campbell, A. W., Cornish, G. B., Martin, D., and Allen, H. M. (2001), "The statistical analysis of quality traits in plant improvement programs with application to the mapping of milling yield in wheat," *Australian Journal of Agricultural Research*, 52, 1207–1219.
- Whitaker, D., Williams, E. R., and John, J. A. (2002), "CycDesigN: A package for the computer generation of experimental designs (Version 2.0)," CSIRO, Canberra, Australia. URL: <http://www.ffp.csiro.au/software>.
- Wilkinson, G. N. and Rogers, C. E. (1973), "Symbolic description of factorial models for analysis of variance," *Applied Statistics*, 22, 392–399.
- Williams, E. R. (1986), "Row and column designs with contiguous replicates," *Australian Journal of Statistics*, 28, 154–163.
- Wood, J. T., Williams, E. R., and Speed, T. P. (1988), "Non-orthogonal block structure in two-phase designs," *Australian Journal of Statistics*, 30A, 225–237.

Table 1. Laboratory phase order and observed Gini coefficients for the biodiversity experiment

| <i>Int3</i> | <i>Int4</i> | <i>Int5</i> | <i>Int6</i> | <i>Run</i> | <i>Method</i> | <i>Block</i> [†] | <i>Plot</i> | <i>Depth</i> | <i>Gini coefficient</i> ($\times 100$) | | | |
|-------------|-------------|-------------|-------------|------------|---------------|---------------------------|-------------|--------------|--|-------|-------|-------|
| | | | | | | | | | <i>Int1</i> <i>Int2</i> 1 | 2 | 1 | 2 |
| 1 | 1 | 1 | 1 | 1 | ground | A | 1 | 0–5 | 66.54 | 66.14 | 65.32 | 63.46 |
| 1 | 1 | 1 | 2 | 2 | ground | A | 1 | 5–10 | 71.45 | 67.24 | 68.64 | 64.34 |
| 1 | 1 | 2 | 1 | 3 | ground | A | 2 | 0–5 | 66.22 | 63.26 | 64.46 | 63.36 |
| 1 | 1 | 2 | 2 | 4 | ground | A | 2 | 5–10 | 67.00 | 63.95 | 68.37 | 63.96 |
| 1 | 2 | 2 | 1 | 5 | ground | B | 1 | 0–5 | 63.90 | 63.53 | 63.91 | 64.11 |
| 1 | 2 | 1 | 2 | 6 | ground | B | 1 | 5–10 | 69.17 | 65.33 | 67.37 | 65.44 |
| 1 | 2 | 1 | 1 | 7 | ground | B | 2 | 0–5 | 64.42 | 61.36 | 63.49 | 62.62 |
| 1 | 2 | 2 | 2 | 8 | ground | B | 2 | 5–10 | 64.02 | 63.36 | 64.84 | 64.03 |
| 2 | 1 | 1 | 1 | 9 | sieved | A | 1 | 0–5 | 66.44 | 69.01 | 66.44 | 68.64 |
| 2 | 1 | 1 | 2 | 10 | sieved | A | 1 | 5–10 | 72.04 | 71.04 | 72.42 | 66.89 |
| 2 | 1 | 2 | 1 | 11 | sieved | A | 2 | 0–5 | 64.90 | 65.72 | 68.24 | 63.81 |
| 2 | 1 | 2 | 2 | 12 | sieved | A | 2 | 5–10 | 70.48 | 70.88 | 71.52 | 70.88 |
| 2 | 2 | 2 | 1 | 13 | sieved | B | 1 | 0–5 | 68.18 | 64.97 | 66.86 | 65.09 |
| 2 | 2 | 1 | 2 | 14 | sieved | B | 1 | 5–10 | 73.05 | 66.89 | 71.86 | 67.74 |
| 2 | 2 | 1 | 1 | 15 | sieved | B | 2 | 0–5 | 65.56 | 62.54 | 64.15 | 65.06 |
| 2 | 2 | 2 | 2 | 16 | sieved | B | 2 | 5–10 | 67.61 | 65.85 | 70.46 | 67.48 |

[†]The letters A and B refer to Blocks 1 and 2, respectively, for the first and third columns of the Gini coefficients and Blocks 3 and 4, respectively, for the other two columns.

Table 2. Analysis of variance for the biodiversity experiment[†]

| <i>Source</i> | <i>df</i> | <i>SSq</i> | <i>MSq</i> | <i>F</i> | <i>p</i> |
|---|-----------|------------|------------|----------|----------|
| Int1 | 1 | 0.72 | 0.72 | 19.98 | 0.140 |
| Int2 [Int1] | 2 | 58.84 | | | |
| Block | 1 | 58.80 | 58.80 | | |
| Residual | 1 | 0.04 | 0.04 | 0.02 | 0.900 |
| Int3 [Int1 \wedge Int2] | 4 | | | | |
| Sample [Block \wedge Plot \wedge Depth] | 2 | 121.82 | | | |
| Method | 1 | 121.28 | 121.28 | | |
| Residual | 1 | 0.54 | 0.54 | | |
| Residual | 2 | 3.19 | 1.59 | 0.43 | 0.677 |
| Int4 [Int1 \wedge Int2 \wedge Int3] | 8 | 64.54 | | | |
| Block | 2 | 37.68 | 18.84 | | |
| Sample [Block \wedge Plot \wedge Depth] | 2 | 11.94 | 5.97 | | |
| Residual | 4 | 14.91 | 3.73 | 2.43 | 0.133 |
| Int5 [Int1 \wedge Int2 \wedge Int3 \wedge Int4] | 16 | 56.72 | | | |
| Plot [Block] | 4 | 43.22 | | | |
| Tillage | 1 | 1.60 | 1.60 | | |
| Residual | 3 | 41.62 | 13.87 | | |
| Sample [Block \wedge Plot \wedge Depth] | 4 | 3.20 | | | |
| Tillage # Method | 1 | 3.04 | 3.04 | | |
| Residual | 3 | 0.16 | 0.05 | | |
| Residual | 8 | 12.30 | 1.54 | 1.58 | 0.207 |
| Int6 [Int1 \wedge Int2 \wedge Int3 \wedge Int4 \wedge Int5] | 32 | 209.67 | | | |
| Depth | 1 | 137.66 | 137.66 | | |
| Block # Depth | 3 | 13.74 | 4.58 | | |
| Plot # Depth [Block] | 4 | 15.87 | | | |
| Tillage # Depth | 1 | 3.87 | 3.87 | | |
| Residual | 3 | 12.00 | 4.00 | | |
| Sample [Block \wedge Plot \wedge Depth] | 8 | 26.86 | | | |
| Depth # Method | 1 | 13.22 | 13.22 | | |
| Tillage # Depth # Method | 1 | 1.01 | 1.01 | | |
| Residual | 6 | 12.63 | 2.10 | | |
| Residual | 16 | 15.54 | 0.97 | | |
| Total | 63 | | | | |

[†]Note that this is an orthogonal analysis so that different instances of the same source must be different orthogonal parts of that source.

Table 3. Laboratory framework for an occasion for the biodiversity experiment

| <i>a) Initial</i> | | | | | | | | <i>b) After blocking</i> | | | | | | | | | | | | | | | |
|-------------------|---|---|---|---|---|---|---|--------------------------|-----------------|---|---|---|---|------------|---|---|---|---|---|---|---|---|--|
| <i>Run</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | <i>Times</i> | 1 | 2 | 3 | 4 | <i>Run</i> | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 2 | |
| <i>Interval</i> | | | | | | | | | <i>Interval</i> | | | | | | | | | | | | | | |
| 1 | | | | | | | | | 1 | | | | | | | | | | | | | | |
| 2 | | | | | | | | | 2 | | | | | | | | | | | | | | |
| 3 | | | | | | | | | 3 | | | | | | | | | | | | | | |
| 4 | | | | | | | | | 4 | | | | | | | | | | | | | | |

Table 4. Decomposition table for the hierarchical design for the biodiversity experiment[†]

| <i>runs tier</i> | | <i>samples tier</i> | | <i>treatments tiers</i> | |
|--|-----------|-----------------------------------|-----------|-------------------------|-----------|
| <i>Source</i> | <i>df</i> | <i>Source</i> | <i>df</i> | <i>Source</i> | <i>df</i> |
| Occasions | 1 | | | | |
| Times [O] | 6 | Blocks | 3 | | |
| | | Residual | 3 | | |
| Intervals \wedge Runs [O \wedge T] | 56 | Depths | 1 | | |
| | | B # D | 3 | | |
| | | Plots [B] | 4 | Tillages | 1 |
| | | | | Residual | 3 |
| | | D # P [B] | 4 | T # D | 1 |
| | | | | Residual | 3 |
| | | Samples [B \wedge P \wedge D] | 16 | Methods | 1 |
| | | | | T # M | 1 |
| | | | | M # D | 1 |
| | | | | T # M # D | 1 |
| | | | | Residual | 12 |
| | | Residual | 28 | | |

[†]The names of sources in decomposition tables follow the convention of using only the first letter of a factor in terms other than the lowest order term in which it occurs.

Table 5. Resolvable incomplete block design for the Beetle damage example

| <i>Rep/ Occasion</i> | I | | | | II | | | | III | | | | IV | | | | V | | | |
|--------------------------|----|----|---|----|----|----|----|----|-----|---|----|----|----|---|----|----|---|----|----|----|
| <i>Block</i> | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| <i>Plate</i> | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| <i>Cell</i> | | | | | | | | | | | | | | | | | | | | |
| 1 | 1 | 2 | 3 | 4 | 1 | 3 | 2 | 4 | 1 | 4 | 2 | 3 | 1 | 2 | 3 | 4 | 1 | 3 | 2 | 4 |
| 2 | 6 | 7 | 8 | 5 | 5 | 7 | 6 | 8 | 8 | 7 | 5 | 6 | 5 | 6 | 7 | 8 | 7 | 5 | 8 | 6 |
| 3 | 11 | 12 | 9 | 10 | 9 | 11 | 10 | 12 | 10 | 9 | 11 | 12 | 12 | 9 | 10 | 11 | 9 | 11 | 10 | 12 |

Table 6. Decomposition table for the beetle damage experiment

| <i>cells tier</i> | | <i>locations tier</i> | | <i>treatments tier</i> | | $E[MSq]^\dagger$ |
|----------------------|-----------|-------------------------------|-----------|------------------------|-----------|---|
| <i>Source</i> | <i>df</i> | <i>Source</i> | <i>df</i> | <i>Source</i> | <i>df</i> | |
| Occasions | 4 | Reps | 4 | | | $\sigma_{OPC}^2 + 3\sigma_{OP}^2 + 12\sigma_O^2 + \sigma_{RBL}^2 + 6\sigma_{RB}^2 + 12\sigma_R^2$ |
| Plates [O] | 15 | Benches [R] | 5 | Damages ¹ | 5 | $\sigma_{OPC}^2 + 3\sigma_{OP}^2 + \sigma_{RBL}^2 + 6\sigma_{RB}^2 + q(D_1)$ |
| | | L ₁ [R \wedge B] | 10 | Damages ² | 8 | $\sigma_{OPC}^2 + 3\sigma_{OP}^2 + \sigma_{RBL}^2 + q(D_2)$ |
| | | | | Residual | 2 | $\sigma_{OPC}^2 + 3\sigma_{OP}^2 + \sigma_{RBL}^2$ |
| Cells [O \wedge P] | 40 | L ₂ [R \wedge B] | 40 | Damages ³ | 11 | $\sigma_{OPC}^2 + \sigma_{RBL}^2 + q(D_3)$ |
| | | | | Residual | 29 | $\sigma_{OPC}^2 + \sigma_{RBL}^2$ |
| Total | 59 | | | | | |

[†]The σ^2 s are variance components with the subscript comprised of the first letter of each factor in the term and $q(D_i)$ is a quadratic function of the expectation of the data whose matrix is the projector D_i for the i th Damages source.

¹canonical efficiency factors: 0.267, 0.267, 0.267, 0.133, 0.067.

²canonical efficiency factors: 0.400, 0.400, 0.333, 0.267, 0.200, 0.133, 0.133, 0.133.

³canonical efficiency factors: 1.000, 1.000, 0.800, 0.800, 0.733, 0.733, 0.600, 0.600, 0.600, 0.600, 0.533.

Table 7. Decomposition table for the nonhierarchical design for the biodiversity experiment

| <i>runs tier</i> | | <i>samples tier</i> | | <i>treatments tier</i> | |
|------------------|-----------|-----------------------------------|-----------|------------------------|-----------|
| <i>Source</i> | <i>df</i> | <i>Source</i> | <i>df</i> | <i>Source</i> | <i>df</i> |
| Occasions | 1 | | | | |
| Runs | 7 | D # P [B] | 1 | | |
| | | Samples [B \wedge P \wedge D] | 2 | T # M # D | 1 |
| | | Residual | 4 | Residual | 1 |
| Intervals [O] | 6 | Blocks | 3 | | |
| | | Residual | 3 | | |
| R # O | 7 | Plots [B] | 1 | | |
| | | B # D | 1 | | |
| | | Samples [B \wedge P \wedge D] | 2 | | |
| | | Residual | 3 | | |
| I # R [O] | 42 | Depths | 1 | | |
| | | B # D | 2 | | |
| | | Plots [B] | 3 | Tillages | 1 |
| | | | | Residual | 2 |
| | | D # P [B] | 3 | T # D | 1 |
| | | | | Residual | 2 |
| | | Samples [B \wedge P \wedge D] | 12 | Methods | 1 |
| | | | | T # M | 1 |
| | | | | M # D | 1 |
| | | | | Residual | 9 |
| Residual | 21 | | | | |

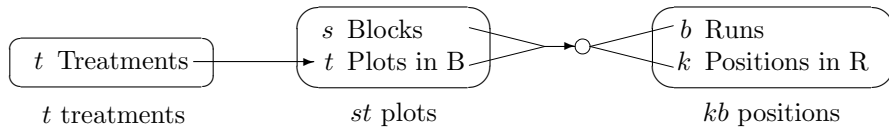


Figure 1. Composed randomizations for a two-phase experiment that uses a balanced incomplete block design

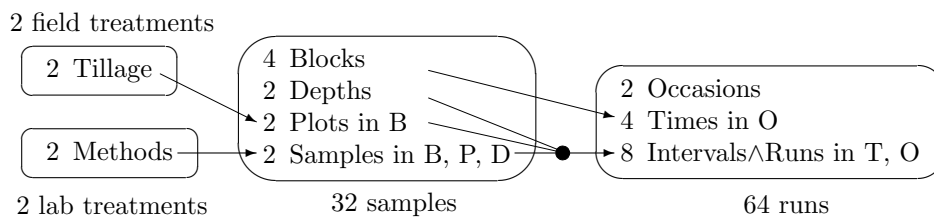


Figure 2. Randomizations for the hierarchical design for the biodiversity experiment.

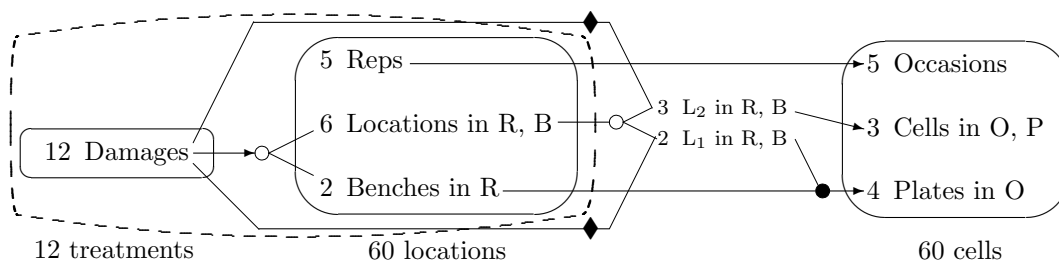


Figure 3. Randomizations for the beetle damage experiment.

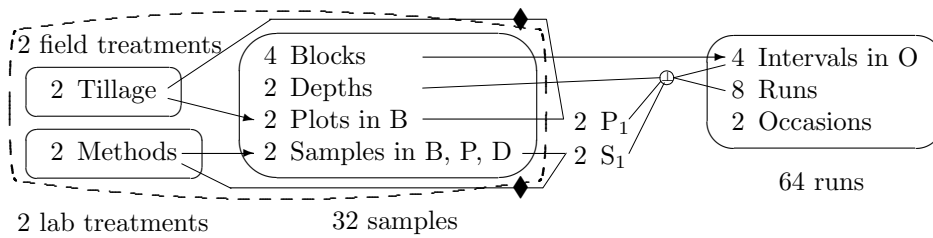


Figure 4. Randomizations for the nonhierarchical design for the biodiversity experiment.