BIOMETRICS INFORMATION
(You're 95% likely to need this information)

PAMPHLET NO. # 55 DATE: November 15, 1996

SUBJECT: Displaying Factor Relationships in Experiments

Determining the proper description of a study requires identifying the independent variables in the study, how they are related to each other, and how they are replicated. The Factor Relationship Diagram (FRD) described here is a useful aid for doing this and for generating the list of sources of variation for an ANOVA table. Once developed for a study, the diagram can be used to discuss replication issues and the effects of any proposed changes. Furthermore, it clarifies the structure of various experimental designs, particularly the often confusing split-plot, and allows the determination of the experimental design for studies that do not follow the traditional patterns presented in most textbooks. The FRD has been used within the B.C. Research Program for many years where it has been commonly called the 'stick diagram'.

This pamphlet will describe how to use the diagram by using a randomized block study example, and then modifying the example to illustrate the factor relations within a split-plot design. Features of these designs, including common pitfalls, are discussed throughout. A review of the basic concepts and terminology used is in Appendix 1 while a formalized description of the steps for diagram construction are presented in Appendix 2.

Randomized Block Design with Subsampling

The construction of a factor relationship diagram will be illustrated with a randomized block design. Suppose that data from the following experiment are to be analyzed. Three blocks of land were each divided into two plots, each of which was randomly assigned one of two levels of a fertilizer treatment, called treatment A. Ten similar nursery seedlings were then planted in each plot and the treatment applied. A year later, the height of each seedling was recorded.

This study has three unit factors[1]: block, plot, and seedling and one treatment factor, A. Each factor is identified by a letter, B, P, S, or A, respectively, and each level of each factor is given a unique number. For example, B has three levels, 1, 2, and 3. The study includes six unique plots, therefore P is assigned numbers 1 through 6. Similarly, the 60 seedlings are uniquely numbered from 1 to 60.

[1] See Appendix 1 for a definition of unit factor.

Blocks, plots, and seedlings are unit factors because seedlings have been planted within plots and plots are grouped within blocks. The diagram for this unit structure is shown in Figure 1. The largest unit factor, blocks, is placed at the top of the diagram, plots are placed below this, and seedlings are at the bottom. Lines are drawn between block 2 and plots 3 and 4, for example, to show that these plots are part of block 2.

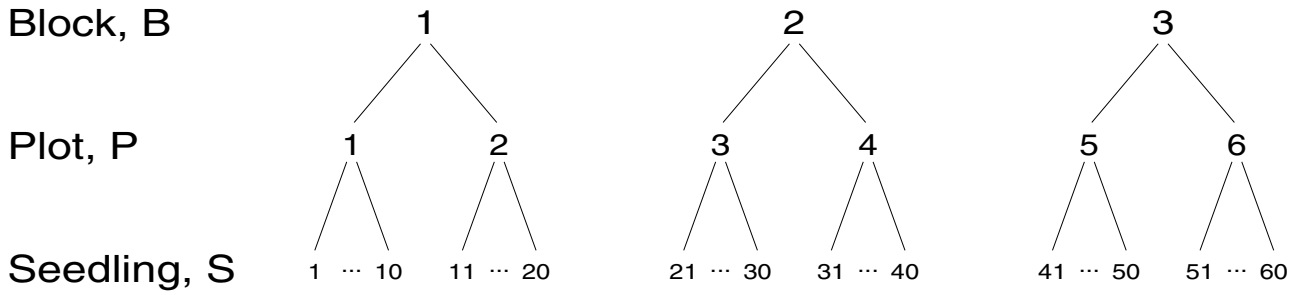

Figure 1. Unit Structure for a randomized block design with subsampling

The treatment structure is quite simple, because it involves only one factor. Because treatments were randomly assigned to the plots, the numbers representing the levels of A are placed in the row above plots. Seedlings are the subsamples or observational units. The final diagram is shown in Figure 2.

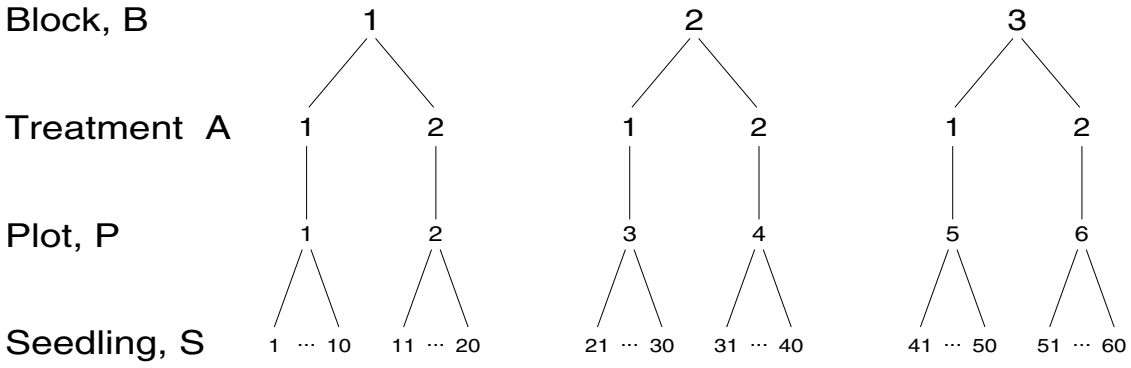

Figure 2. First Factor Relationship Diagram for the Example: A Randomized Block Design.

It is readily apparent from this diagram that factors B and A are crossed, since each level of A occurs with each level of B. However, P is nested within each combined level of B and A, denoted by P(AB), since each plot can only be located within one block and be assigned one level of A. Similarly, each seedling is nested within each plot within each combination of B and A, denoted by S(PAB). The sources of variation for this design are shown in Table 1. The expected mean squares for all the ANOVA tables are based on the assumption that treatment is a fixed factor while block, plot, and seedling are random factors (some references on calculating expected mean square are listed at the end of Appendix 2).

Table 1. The analysis of variance for the design shown in figure 2

Sources of variation		*Degrees of freedom*	*Expected mean squares*
Block	B	b-1 = 2	$\sigma^2_{S(PAB)} + 10\ \sigma^2_{P(AB)} + 20\ \sigma^2_{B}$
Treatment	A	a-1 = 1	$\sigma^2_{S(PAB)} + 10\ \sigma^2_{P(AB)} + 10\ \sigma^2_{AB} + 30\ \phi_A$
A x B	A x B	(a-1)(b-1) = 2	$\sigma^2_{S(PAB)} + 10\ \sigma^2_{P(AB)} + 10\ \sigma^2_{AB}$
Plots	P(AB)	ab(p-1) = 0	$\sigma^2_{S(PAB)} + 10\ \sigma^2_{P(AB)}$
Seedlings	S(PAB)	pab(s-1) = 54	$\sigma^2_{S(PAB)}$
Total		pabs-1 = 59	

The degrees of freedom corresponding to each source depends on the number of levels for factors B, P, S, and A, which are denoted by the lower case letters b, p, s, and a, respectively. For this example, b = 3 and a = 2. Because there is only one plot per A x B combination p = 1, even though there are six plots in all. Similarly, s = 10 because each plot contains 10 seedlings. The resulting degrees of freedom for the ANOVA are shown in Table 1. The total degrees of freedom (the total number of observations minus one) is included as a useful check of the calculations.

Because there are subsamples within each plot, this randomized block design is different from most textbook examples (although see Anderson and McLean 1974 sec. 5.3; Hinkelmann and Kempthorne 1994, pg. 182-187; Kempthorne 1952 pp. 212-216; Kuehl 1994 pp. 15-17 and sec. 5.9; and Steel and Torrie 1980 sec. 9.8). Most discussions of the randomized block design do not comment on the role of the plots or parts of the block to which the treatments are assigned, and it is easy to assume that seedlings are the experimental units for the treatment A. Furthermore, lines with zero degrees of freedom and their associated variance components are usually omitted from the ANOVA table. Their inclusion is helpful, however, when learning about replication and sources of variation. For instance, while $\sigma^2_{P(AB)}$ is not estimable with this design (since it has zero

degrees of freedom), it is an important component in the Expected Means Squares for A. The diagram in Figure 2 and resulting ANOVA in Table 1 clearly show that plots are the experimental units for A, while seedlings are the subsamples.

Blocks are a random factor because they are assumed to have been randomly selected from a larger collection of blocks, thus allowing the results to be generalized to this larger collection. The resulting expected mean squares in Table 1 show that the error term for A is A x B. This means that the A effect must be larger than the A x B interaction, which is a measure of the consistency of the main effects of A across the blocks. Using this to test A makes intuitive sense because inferences about the behavior of A to blocks not in the study is not reasonable if the main effects of A are not consistent across those in the study. While it is often assumed that σ^2_{AB} is zero, this assumption is unnecessary, because the A main effect can be larger than its interaction with block. This will occur when treatment rankings are similar for each block, even though the differences between pairs of treatments change substantially from block to block.

Figure 2 and Table 1 provide a focus for discussion of treatment replication. First, it is clear that plots, not seedlings, are the experimental units for A, even though plots do not directly form the error term for A. Second, blocks provide the replication, and their number determines the degrees of freedom for the error term. Thus, to increase replication, it is clear that increasing the number of blocks is preferable to increasing p, the number of plots per treatment per block. If p is increased then P(AB) no longer has zero degrees of freedom and the resulting design is known as the generalized randomized block design (Steel and Torrie 1980, p. 215).

Split-plot Design: Non-replicated

To continue with the example, suppose that during discussion of the results it became clear that differences between the blocks are of particular interest. This might be because the blocks were chosen from areas of different environmental conditions; for example, one block might be xeric, another mesic, while the third is hygric. Hence, a new treatment, moisture regime, must be identified and conceptually separated from its experimental unit, the block. While it is good experimental design to select blocks that differ from each other (while being homogeneous within), inferences about differences between them are not appropriate. The corresponding change to the design is facilitated by examining the original factor relationship diagram (Figure 2), and adding a new row at the top for moisture regime (M). The resulting diagram is shown in Figure 3.

This experiment has become a split-plot with no replication of the main-plot factor (M). The new ANOVA is shown in Table 2. This table has several sources of variation with zero degrees of freedom. Notice that the error term for M is B(M), but the corresponding variance component ($\sigma^2_{B(M)}$) is not estimable. Furthermore, the tests for A and A x M require the variance component

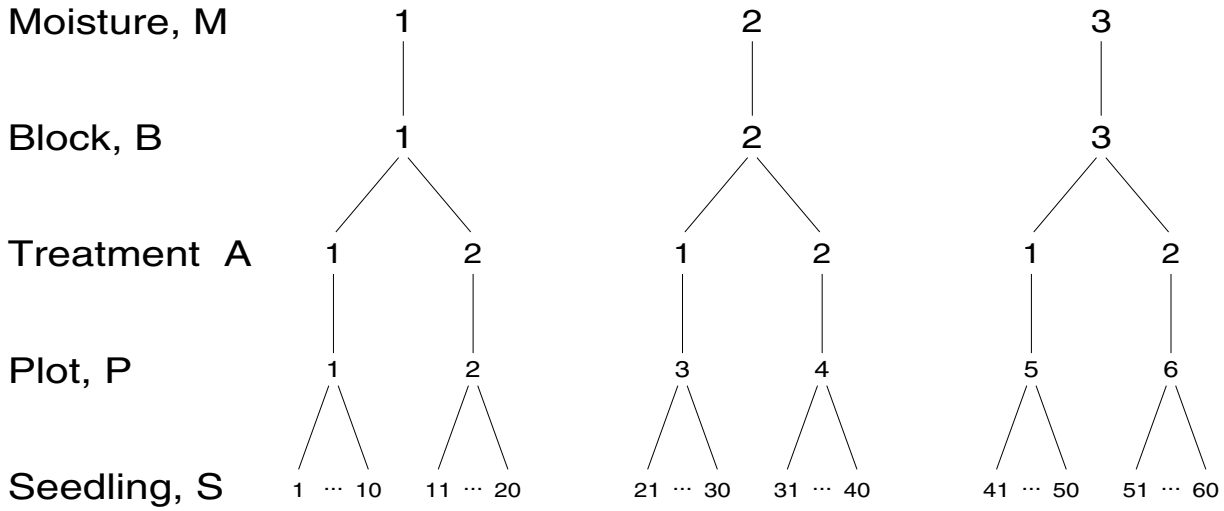

Figure 3. Factor Relationship Diagram for the Example with Moisture Added: A Split-plot Design.

Table 2. The analysis of variance for the design shown in figure 3

Sources of variation		*Degrees of freedom*	*Expected mean squares*
Moisture	M	m-1 = 2	$\sigma^2_{S(PABM)} + 10\ \sigma^2_{P(ABM)} + 20\ \sigma^2_{B(M)} + 20\ \phi_M$
Block	B(M)	m(b-1) = 0	$\sigma^2_{S(PABM)} + 10\ \sigma^2_{P(ABM)} + 20\ \sigma^2_{B(M)}$
Treatment	A	a-1 = 1	$\sigma^2_{S(PABM)} + 10\ \sigma^2_{P(ABM)} + 10\ \sigma^2_{AB(M)} + 30\ \phi_A$
A x M	A x M	(a-1)(m-1) = 2	$\sigma^2_{S(PABM)} + 10\ \sigma^2_{P(ABM)} + 10\ \sigma^2_{AB(M)} + 10\ \phi_{AM}$
A x B(M)	A x B(M)	m(a-1)(b-1) = 0	$\sigma^2_{S(PABM)} + 10\ \sigma^2_{P(ABM)} + 10\ \sigma^2_{AB(M)}$
Plots	P(ABM)	abm(p-1) = 0	$\sigma^2_{S(PABM)} + 10\ \sigma^2_{P(ABM)}$
Seedlings	S(PABM)	pabm(s-1) = 54	$\sigma^2_{S(PABM)}$
Total		pabms-1 = 59	

$\sigma^2_{AB(M)}$, which is also not estimable. If it is reasonable to assume that σ^2_{AM} is zero, then A x M could be used to test A.

On the other hand, it is easy to not recognize the proper experimental units for the treatments and improperly identify this as a simple two-way factorial, with seedlings as the experimental units for both treatments. The resulting ANOVA is shown in Table 3. This is very misleading because it implies a factor relationship diagram similar to the one in Figure 4, which is quite different from

that of Figure 3. Moisture regime and blocks are confounded and, in this ANOVA, all differences between them have been ascribed to moisture (notice where the block's degrees of freedom have gone in the move from Table 1 to Table 3). Tests of the factors M, A, and A x M are now made with the seedling variability. The Expected Means Squares in Table 2 show that the new F-tests will be biased in favor of detecting treatment differences, because it is unlikely that $\sigma^2_{B(M)}$, $\sigma^2_{AB(M)}$, and $\sigma^2_{P(ABM)}$ are all zero.

Table 3. The analysis of variance table as it might erroneously appear in a final report

Sources of variation		*Degrees of freedom*	*Expected mean squares*
Moisture	M	m-1 = 2	$\sigma^2_{S(AM)} + 20\ \phi_M$
Treatment	A	a-1 = 1	$\sigma^2_{S(AM)} + 30\ \phi_A$
A x M	A x M	(a-1)(m-1) = 2	$\sigma^2_{S(AM)} + 10\ \phi_{AM}$
Seedlings	S(AM)	am(s-1) = 54	$\sigma^2_{S(AM)}$
Total		ams-1 = 59	

If A x M cannot be assumed negligible, then the analysis of Table 3 may be all that is available. In this case, a final report should include a discussion of the confounded effects and justify the assumption of homogeneity of the experimental material on non-statistical grounds. Any conclusions from this study would be tentative, exploratory in nature and require confirmation with a follow-up experiment. Designs such as this arise often and are too frequently analyzed using ANOVA's such as those in Table 3 without proper acknowledgment of their weaknesses.

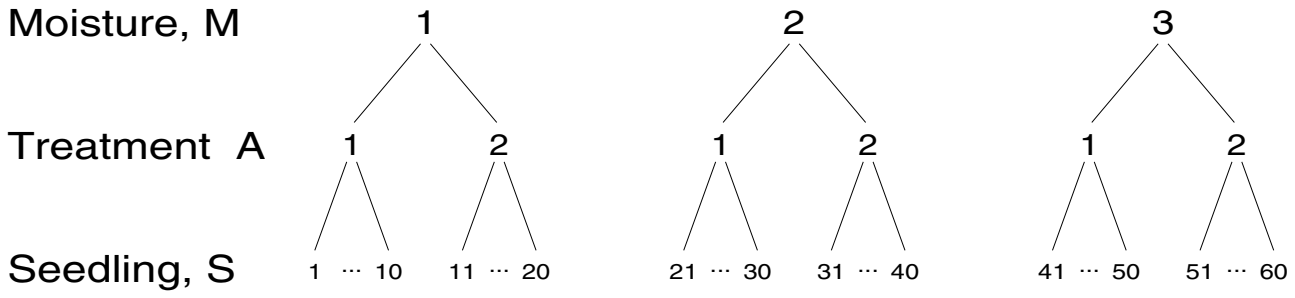

Figure 4. Factor Relationship Diagram for the Simple Factorial Implied by Table 3.

Split-plot Design: Replicated

To extend the split-plot discussion further, suppose that a follow-up study is planned or that this discussion occurred during the planning stages of the experiment. Since moisture regime is a classification factor of interest, it should be properly replicated. If blocks are the appropriate experimental unit, then at least two blocks within each moisture regime should be included. The corresponding factor relationship diagram is shown in Figure 5 and the ANOVA in Table 4. Note that this is a split-plot design based on the completely randomized design for the main-plot factor, which is different from a split-plot based on the randomized block design for the main-plot factor (see Milliken and Johnson 1984, pp. 73, 76).

The distinguishing feature between the split-plot and the simple factorial designs is that, while both have at least two treatment/classification factors with a crossed arrangement, in the split-plot these factors are applied to at least two different experimental units, which have a nested relationship (and are of different sizes). The difference in these relationships can be appreciated by comparing Figures 4 and 5 (although the plots are omitted from Figure 4).

Table 4. The analysis of variance for the design shown in figure 5

Sources of variation		*Degrees of freedom*	*Expected mean squares*[a]
Moisture	M	m-1 = 2	$\sigma^2_{S(PABM)} + 10\,\sigma^2_{P(ABM)} + 20\,\sigma^2_{B(M)} + 40\,\phi_M$
Block	B(M)	m(b-1) = 3	$\sigma^2_{S(PABM)} + 10\,\sigma^2_{P(ABM)} + 20\,\sigma^2_{B(M)}$
Treatment	A	a-1 = 1	$\sigma^2_{S(PABM)} + 10\,\sigma^2_{P(ABM)} + 10\,\sigma^2_{AB(M)} + 30\,\phi_A$
A x M	A x M	(a-1)(m-1) = 2	$\sigma^2_{S(PABM)} + 10\,\sigma^2_{P(ABM)} + 10\,\sigma^2_{AB(M)} + 10\,\phi_{AM}$
A x B(M)	A x B(M)	m(a-1)(b-1) = 3	$\sigma^2_{S(PABM)} + 10\,\sigma^2_{P(ABM)} + 10\,\sigma^2_{AB(M)}$
Plots	P(ABM)	abm(p-1) = 0	$\sigma^2_{S(PABM)} + 10\,\sigma^2_{P(ABM)}$
Seedlings	S(PABM)	pabm(s-1) = 108	$\sigma^2_{S(PABM)}$
Total		pabms-1 = 119	

[a] The expected mean squares are based on the assumption that treatment and moisture are fixed factors.

Discussion

The exercise of drawing the factor relationship diagram helps us understand the relationships between various factors in an experiment, and often forces the proper identification of each factor's experimental unit. This is very helpful when trying to understand randomized block and split-plot designs and can demonstrate that what appeared to be a simple factorial could be a split-plot with

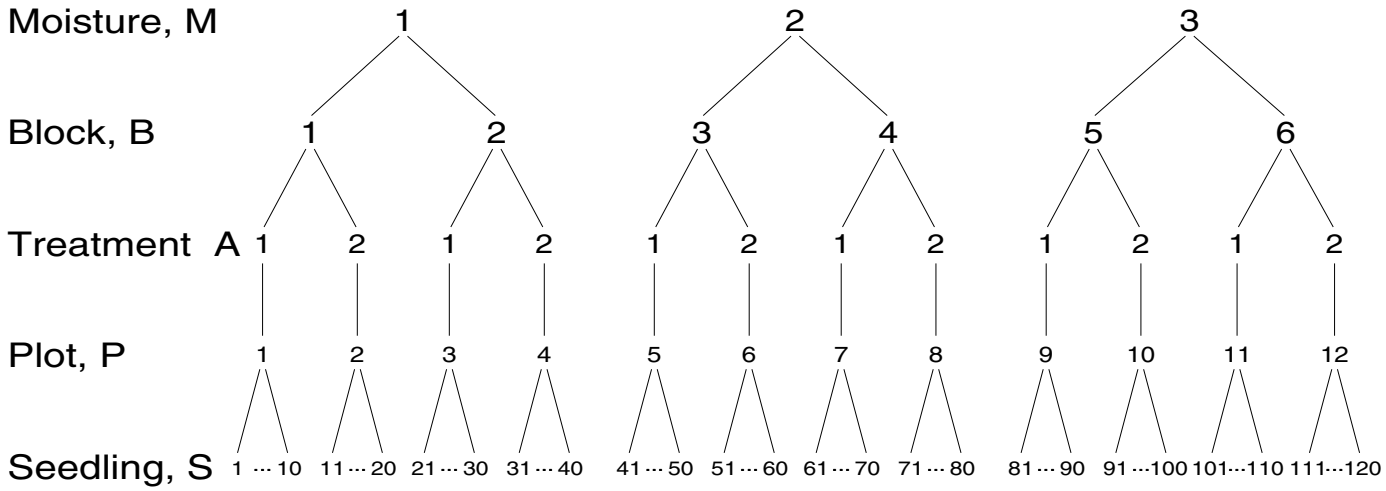

Figure 5. Factor Relationship Diagram for the Full Split-plot Design.

no proper replication (as in the second example). This ability to identify the proper experimental units, even if they have zero degrees of freedom, helps obtain a correct model for analysis, helps decide if treatments are replicated and facilitates discussions about methods of replication. When treatments cannot be replicated (e.g., when only two greenhouses are available to create different artificial daylengths), the diagram is very useful in determining the complete ANOVA table and identifying what assumptions are necessary for the analysis. The scientist can then argue for the reasonableness of these assumptions on non-statistical grounds.

When studies have only one experimental unit assigned to each treatment, the lack of replication can be obscured by the occurrence of several subsamples within each experimental unit. Failing to separate the subsamples from the experimental units and analyzing the data from the subsamples as if each were a proper experimental unit greatly increases the apparent replication. This mistake has been called ***pseudo-replication*** (Hurlbert 1984) and is more easily avoided if the design of the experiment is carefully diagramed.

A simple method for analyzing designs with subsamples is to calculate experimental unit means from the subsamples and to use those means in the analysis (Steel and Torrie 1980, p. 216). A second method is to include them in the ANOVA table, as has been discussed here. The factor relationship diagram helps ensure that they are placed correctly in the experimental design and are not inappropriately used as error terms.

The process of drawing the factor relationship diagram helps elucidate the structure of all the usual balanced complete designs and their many variations, which are rarely discussed in textbooks. When designing experiments, the factor relationship diagram can be used to describe likely designs, determine their ANOVA tables, and then compare their relative advantages and weaknesses. The factor relationship diagram helps match the statistics to the logic of the experiment, instead of adapting the experiment to the closest familiar statistical recipe.

Prepared by:

Wendy A. Bergerud, Research Branch, B.C. Ministry of Forests, Victoria, B.C. V8W 9C2.

[This material has been adapted from a paper published in American Statistician, August 1996, Vol 50, No 3, pp228-233]

Bibliography

Anderson, V. L. and McLean, R. A. (1974a), *Design of Experiments*: *A Realistic Approach*, New York: Marcel Dekker, Inc.

Cox, D. R. (1958), *Planning of Experiments*, New York: John Wiley.

Hicks, C. R. (1973), *Fundamental Concepts in the Design of Experiments*, (2nd ed.), New York: Holt, Rinehart, and Winston.

Hinkelmann, K. and Kempthorne, O. (1994), *Design and Analysis of Experiments: Volume I Introduction to Experimental Design*, John Wiley, New York.

Hurlburt, S. H. (1984), "Pseudo-replication and the design of ecological field experiments," *Ecological Monographs*, 54, 187-211.

Kempthorne, O. (1952), *Design and Analysis of Experiments*, Malabar, FL: Robert E. Krieger Publishing.

Keppel, G. (1982), *Design and Analysis*: *A Researcher's Handbook* (2nd ed.), Englewood Cliffs, NJ: Prentice-Hall.

Kuehl, R. O. (1994), *Statistical Principles of Research Design and Analysis*, Belmont, CA: Duxbury Press.

McPherson, G. (1990), *Statistics in Scientific Investigation*: *Its Basis*, *Application*, *and Interpretation*, New York: Springer-Verlag.

Mason, R. L., Gunst, R. F., and Hess, J. L. (1989), *Statistical Design and Analysis of Experiments*: *with Applications to Engineering and Science*, New York: John Wiley.

Maxwell S. E., and Delaney H. D. (1990), *Designing Experiments and Analyzing Data*: *A Model Comparison Perspective*, Belmont, CA: Wadsworth Publishing.

Mead, R. (1988), *The Design of Experiments*: *Statistical principles for practical application*. Cambridge, U.K.: Cambridge University Press.

Milliken, G. A., and Johnson, D. E. (1984), *Analysis of Messy Data*, *Volume I*: *Designed Experiments*, Belmont, CA: Lifetime Learning Publications.

——————— (1989), *Analysis of Messy Data*, *Volume II*: *Nonreplicated Experiments*, New York: Van Nostrand Reinhold.

Schwarz, C. J. (1993), "The Mixed-Model ANOVA: The Truth, the Computer Packages, the Books (Part I: Balanced Data)," *The American Statistician*, 47, 48-59.

Sit, V. 1995. Analyzing ANOVA Designs. Biom. Info. Hand. 5, Res. Br., B.C. Min. For., Victoria, B.C. Work. Pap. 07/1995.

Steel, R. G. D., and Torrie, J. H. (1980) *Principles and Procedures of Statistics*: *A Biometrical Approach* (2nd ed.), New York: McGraw-Hill.
Taylor, W. H., and Hilton, H. G. (1981), "A Structure Diagram Symbolization for Analysis of Variance," *The American Statistician*, 35, 85-93.
Winer, B. J. (1971), *Statistical Principles in Experimental Design* (2nd ed.), New York: McGraw-Hill.

Appendix 1: Basic Concepts

Factors and Their Relationships

A ***factor*** is an independent variable that may affect the response variable. A ***treatment factor*** is an independent variable that the researcher controls in the study. Examples of treatment factors include site preparation, herbicide application, or amounts of fertilizer. A ***classification factor*** is an independent variable that cannot be assigned by the researcher (Cox 1958 pp. 92, 135; Keppel 1982, p. 8), but is either of direct interest or needs to be accounted for within the study. Examples of classification factors include subzone, gender, or tree species, which are intrinsic to the experimental material in hand. The values that the factors can or do take on are known as ***levels***. Gender, for instance, has two common levels: male and female. Tree species levels might include Douglas fir, Western Red Cedar and Birch. Factors may have a ***crossed*** relationship if each level of one factor occurs with each level of the other, or one factor may be ***nested*** within another if its levels occur with only one level of the other factor. The treatment and classification factors, together with their relationships, form the ***treatment structure*** of an experiment.

Experimental Units and Their Structure

Experimental units are those basic units of experimental material to which a treatment level will be applied, and each should have an equal chance of receiving any of the treatments (random assignment). This definition must be modified somewhat for classification factors, because experimental units cannot be randomly assigned a level of the factor. For example, trees can have their species identified, but this attribute cannot be randomly assigned. Thus a better definition of experimental units for classification factors is that basic unit of experimental material to which a classification level can be identified. Consider plants that produce both male and female flowers, so that an individual plant cannot be identified as either male or female. In this case, flowers are the appropriate experimental units for gender. All experimental units should be randomly sampled from a well-defined population of such units to which the study results will be generalized (random selection).

Many studies contain experimental units that are either subsampled or contain two or more elements. For instance, plots of land may each have five soil pits and experimental units might be

rows of 20 trees. Subsampling occurs in these instances because the response variable(s) must be measured separately for each soil pit or tree. These subsamples or elements have also been called ***observation units*** (Kuehl 1994) and ***sampling units*** (Steel and Torrie 1980 p. 124).

Experimental units occur at one or more levels of a hierarchical structure of the experimental material. This structure is independent of the treatment structure and describes, for example, the grouping of experimental units into blocks, and any subsampling or repeated measures of the experimental units. These factors are not of direct interest in themselves but are necessary to describe the methods of measurement, replication, and randomization. Their levels may be an intrinsic property of the experimental material (e.g., when orchard trees have several treatments applied to different branches) or they may be controlled by the researcher (e.g., when potted trees are grouped onto greenhouse benches). This structure is called the ***unit structure*** by Mead (1988, p. 481) and the ***design structure*** by Milliken and Johnson (1984, sec. 4.2). Factors involved in this structure, such as main-plots, seedlots, plots, or stands, will be called ***unit factors***.

Experimental Design

The ***experimental design*** is obtained by describing the assignment of treatments and classification factors to their respective experimental units within the unit or design structure. Randomized block and split-plot designs are common examples of experimental designs.

Appendix 2: Rules for Diagram Construction

The steps for construction of the factor relationship diagram are:

1. The treatment, classification, and unit factors are identified. Unique levels of each factor are assigned unique numbers (although these same numbers may be used for other factors). This method of assignment is essential because it allows easy identification of factor relationships (i.e., whether nested or crossed).
2. The unit structure is drawn with the largest grouping of experimental material placed on the top row and successively smaller groupings placed below. Lines can be drawn between rows to show which levels of each factor occur with each other (they may need to redrawn in the next step). Unit factors are nested within larger unit factors but can be nested or crossed with treatment and classification factors.
3. The method of assignment of each treatment or identification of each classification factor to the experimental material is closely examined to determine which unit factor is the actual experimental unit. Each factor is then placed in the unit structure on a row **above** its experimental units, and lines between rows are drawn to indicate which experimental units receive which treatment (the actual randomization is omitted from the diagram to assist

identification of the factor relationships). If two or more factors have the same experimental units, they are placed above that unit factor in any order. Iteration between steps 2 and 3 is sometimes necessary as the experimental design becomes more fully understood.

The sources of variation for an ANOVA can be obtained from a factor relationship diagram by first naming each row or factor. Nested factors are given a complete designation to indicate the hierarchy of levels (e.g., P(AB) for factor P nested within both factors A and B), while all other factors are assigned stand-alone letters (e.g., A and B). A list of sources for the ANOVA table can be generated by combining this set of names in every possible way and eliminating any with the same letter both inside and outside any nesting brackets (see Keppel 1982, pp. 635-637; Winer 1971, sec. 5.14).

This list of sources can be used to specify appropriate logistic regression, log-linear, or ANOVA models. The degrees of freedom and the accompanying expected means squares for balanced complete ANOVA designs can be obtained by following rules found in *Biometrics Handbook No. 5*: *Analyzing ANOVA Designs*. Other sources include Winer (1971), Hicks (1973, pp. 177-182), Anderson and McLean (1974, pp. 52-56), Taylor and Hilton (1981), Keppel (1982, pp. 635-664), Milliken and Johnson (1984, chap. 18), Mason, Gunst, and Hess (1989, pp. 371-373), and Kuehl (1994, pp. 246-250). These rules require the additional step of specifying which factors are random and which are fixed and are somewhat controversial for mixed model ANOVA's; see Schwarz (1993) for a recent discussion.