Evaluating And Interpreting Interactions

By

Klaus Hinkelmann

December 13, 2004

Technical Report Number 04-5

DEPARTMENT OF STATISTICS
VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
BLACKSBURG, VIRGINIA 24061

EVALUATING AND INTERPRETING INTERACTION

Klaus Hinkelmann
Department of Statistics
Virginia Tech
Blacksburg, VA 24061

1 INTRODUCTION AND SUMMARY

The notion of interaction plays an important − and sometimes frightening − role in the analysis and interpretation of results from observational and experimental studies. In general, results are much easier to explain and to implement if interaction effects are not present. It is for this reason that they are often assumed to be negligible. This may, however, lead to erroneous conclusions and poor actions.

One reason why interactions are sometimes feared is because of limited understanding of what the word "interaction" actually means, in a practical sense and,in particular, in a statistical sense. As far as the latter is concerned, simply stating that interaction is significant is generally not sufficient. Subsequent interpretation of that finding is needed, and that brings us back to the definition and meaning of interaction within the context of the experimental setting.

In the following sections we shall define and discuss various types of variables that affect the response and the types of interactions among them. These notions will be illustrated for one particular experiment to which we shall return throughout our discussion. To help us in the interpretation of interactions we take a closer look at the definitions of two-factor and three-factor interactions in terms of simple effects. This is followed by a discussion of the nature of interactions and the role they play in the context of the experiment, from the statistical point of view and with regard to the interpretation of the results. After a general overview of how to dissect interactions we return to our example and perform a detailed analysis and interpretation of the data using SAS® (SAS Institute, 2000), in particular PROC GLM and some of its options, such as SLICE. We mention also different methods for the analysis when interaction is actually present. We conclude the analytical part with a discussion of a useful graphical method when no error term is available for testing for interactions. Finally, we summarize the results with some recommendation reminding the reader that in all of this the experimental design is of fundamental importance.

2 TYPES OF VARIABLES AND INTERACTIONS

The individual response in an observational or experimental study can be described in its most general form as

$$\text{Response} = f(\text{Explanatory Variables}) + \text{error} \tag{1}$$

where f in **(1)** represents a generally unknown function. For practical purposes and as a first approximation, however, we consider most often f as a linear function. Furthermore, it is useful to subdivide the explanatory variables in **(1)** into three categories and rewrite **(1)** as

$$Y = f(x_1,\, x_2,\, \ldots,\, x_t;\, z_1,\, z_2,\, \ldots,\, z_q; u_1,\, u_2,\, \ldots,\, u_s) + e \tag{2}$$

where Y represents the response and, following Cox (1984), $\{x_1,\, x_2,\, \ldots,\, x_t\} \equiv X$ represent treatment variables, $\{z_1,\, z_2,\, \ldots,\, z_q\} \equiv Z$ represent intrinsic variables, $\{u_1,\, u_2,\, \ldots,\, u_s\} \equiv U$ represent nonspecific variables, and e represents an error.

Among these three categories the treatment variables (X) are the major focus in any investigation. In an experimental study they represent different levels of treatment factors which are randomly assigned to the experimental units. In an observational study they represent different values of such explanatory variables that in an experimental study would have been under the control of the investigator. The levels of treatment variables may be binary, nominal, ordinary or quantitative. Some examples are: Type of fertilizer, amount of fertilizer, type of soil preparation, pH level, planting date, plant density, pruning method, irrigation method. water level, variety.

The intrinsic variables (Z) are often introduced by the investigator in order to broaden the scope of the investigation, i.e., statistical inference. For example, an investigation using only one variety of a certain plant may not be particularly useful, but using several, often carefully selected, diverse varieties of this plant will assure a broader applicability of the results of the study. In this sense the intrinsic variables represent blocking factors, the levels of which may be binary, nominal, ordinal or quantitative. Other examples, some representing environmental conditions, are: Species, age of plants, soil type, pH level, humidity, time.

The reader will have noticed that some factors, e.g. variety, appear as both treatment and intrinsic factors. The distinction reflects the fact whether the factor, or rather its levels, are being assigned randomly to experimental units or whether they define, partially or completely, an experimental unit. This is important from a statistical point of view since significance tests can be performed only about treatment factors but not about blocking factors (see Hinkelmann and Kempthorne, 1994, Chapter 9).

The non-specific variables (U) represent typically blocking factors which are being introduced for purely experimental and statistical reason. Their main purpose is to reduce experimental variability and thereby improve statistical power of the tests for treatment factors (see Hinkelmann and Kempthorne, 1994, Chapter 9). Some examples are: Replicates, blocks, families of plants, pieces of land, benches in a greenhouse. Their levels are usually nominal.

With this classification in mind we can be more specific in describing the interactions we are mostly interested in evaluating and explaining, both from a statistical and subject-matter point of view. For two-factor interactions these can be described symbolically as $X \times X$, $X \times Z$, and $X \times U$. Similarly, for three-factor interactions our interest is in $X \times X \times X$, $X \times X \times Z$, and $X \times X \times U$. This list can be extended, of course, but typically higher-order interactions are negligible, or negligible from a practical point of view.

3 AN EXAMPLE

The following example, which was discussed by Pearce (1953, 1983), will be used to illustrate the concepts mentioned in Section 2.

The objective is to study the effect of different pruning methods on the yield of varieties of pears. There are two treatment factors: $x_1 \equiv A$ = type of pruning, $x_2 \equiv B$ = amount of pruning, each with two levels. For factor A the two levels are: F = pruning with few leaders, M = pruning with many leaders, and for factor B the two levels are: H = hard pruning, L = light pruning. In order to broaden the scope of the study, the investigator included five varieties of pears: Am = Beurré d'Amanlis, Ha = Beurré Hardy, Co = Conference, Fe = Fertility, Pi = Pitmaston. These constitute the five levels of the intrinsic factor $z_1 = V$ = variety. The experiment was set up as a randomized complete block design with six blocks for each variety (see Exhibit 1). Thus there is one non-specific factor $u_1 \equiv \beta$ = block with six levels (the original experiment had eight blocks for each variety).

Thus, in summary, the experiment is a 2^2 factorial experiment with treatments (F, H), (M, H), (F, L), (M, L) in a randomized complete block design with a factorial blocking structure ($V \times \beta$) and $5 \times 6 = 30$ blocks of size four each. The four treatments were randomly assigned to four experimental units (trees) in each block.

Denoting the response to the treatment by Y, we can write out a linear model analog to **(2)** reflecting the treatment and block structures and the type of interactions mentioned in Section 2, as follows:

$$\begin{aligned} Y_{ijk\ell} = \mu \quad & + \quad V_i + \beta_{ij} + A_k + B_\ell + (AB)_{k\ell} \\ & + \quad (VA)_{ik} + (VB)_{i\ell} + (VAB)_{ik\ell} \\ & + \quad (\beta A)_{ijk} + (\beta B)_{ij\ell} + (\beta AB)_{ijk\ell} + e_{ijk\ell} \end{aligned} \tag{3}$$

where

EXHIBIT 1

EXPERIMENTAL LAYOUT (SCHEMATIC):

	Block 1	2	3	4	5	6
Am	(L, F) (H, M (H, F) (L, M)	(H, M) (L, M) (L, F) (H, F)				
Ha						
Co						
Fe						
Pi						

μ	=	mean,
V_i	=	effect of i-th variety ($i = 1, 2, \ldots, 5$),
β_{ij}	=	effect of j-th block in i-th variety ($j = 1, 2, \ldots, 6$),
A_k	=	effect of k-th type of pruning ($k = 1(F), 2(M)$),
B_ℓ	=	effect of ℓ-th amount of pruning ($\ell = 1(H), 2(L)$),
$(AB)_{k\ell}$	=	$X \times X$ interaction,
$(VA)_{ik}$, $(VB)_{i\ell}$	=	$X \times Z$ interaction,
$(VAB)_{ik\ell}$	=	$X \times X \times Z$ interaction,
$(\beta A)_{ijk}$, $(\beta B)_{ij\ell}$	=	$X \times U$ interaction,
$(\beta AB)_{ijk\ell}$	=	$X \times X \times U$ interaction,
$e_{ijk\ell}$	=	error associated with the experimental unit in i which received the treatment combination $A_k B_\ell$.

Based on model **(3)** we can partition the total number of degrees of freedom, $119 = 120 - 1$, in the analysis of variance (ANOVA) table as given in Table 3. We note here that the effect terms contained in model **(3)** account for all the degrees of freedom (d.f.). As a consequence, no d.f. remain for error. This may appear strange, but we shall return to this point in Section 7. In fact, we encourage the reader to always write out a model such that the effects and interactions account for the total number of d.f. This will provide a check whether in particular all interactions have been accounted for and what assumptions would have to be made to obtain an adequate number of d.f. for error (in addition to possibly existing d.f. for pure error provided by certain experimental designs, such as generalized randomized block design).

TABLE 1: ANOVA FOR MODEL **(3)**

Source of Variation	Degrees of Freedom
V	4
β	$25 = 5(6-1)$
A	1
B	1
$A \times B$	1
$V \times A$	4
$V \times B$	4
$V \times A \times B$	4
$\beta \times A$	25
$\beta \times B$	25
$\beta \times A \times B$	25
Sub-total	119
Error	0
Total	119

4 DEFINING INTERACTIONS

In order to analyze and interpret the types of interactions we have alluded to earlier, it is important to understand how interactions are defined. In this section we shall provide some general definitions and then show how they can be made more concrete in the case of a factorial structure.

4.1 Two-factor interactions

We consider first interactions of the type $X \times X$. Suppose that the function f in **(2)** is of the form $f(x_1, x_2)$, i.e., a function of two treatment factors only. If this function can be expressed as

$$f(x_1, x_2) = f_1(x_1) + f_2(x_2) \tag{4}$$

then there is no interaction between x_1 and x_2. This implies that a change in the response due to a change in x_1 does not depend on the level of x_2, and a change in the response due to a change in x_2 does not depend on the level of x_1.

As an example of **(4)** we mention the familiar multiple linear regression expression (with x_1 and x_2 representing two quantitative treatment variables)

$$f(x_1, x_2) = \alpha + \beta_1 x_1 + \beta_2 x_2 \tag{5}$$

where an increase in x_1 (or x_2) by one unit leads to a change in the response of β_1 (or β_2) units regardless of the level of x_2 (or x_1).

To the extent that **(4)** is not satisfied is an indication of interaction between the two factors. In that case a change in the response due to a change in x_1 (x_2) does depend on the level of x_2 (x_1). As an example consider

$$f(x_1, x_2) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \tag{6}$$

Now a change in x_1 by one unit leads to a change in the response by $\beta_1 + \beta_{12} x_2$ units. Thus the change is no longer independent of the level of x_2.

We now consider the case where the treatment factors exhibit a factorial structure. For simplicity we again consider two factors $x_1 = A$ and $x_2 = B$ say, with two levels each, a_0, a_1 and b_0, b_1, respectively, so that the treatment combinations are a_0b_0, a_1b_0, a_0b_1, a_1b_1. One way of writing out a no-interaction model, corresponding to **(4)** and **(5)**, is

$$f(a_i, b_j) = \mu + \alpha_i + \beta_j \tag{7}$$

and a model with interaction, corresponding to **(6)**, is

$$f(a_i, b_j) = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} \tag{8}$$

where $i = 0, 1$; $j = 0, 1$; α_i represents the effect of factor A at level i; β_j represents the effect of factor B at level j; and $(\alpha\beta)_{ij}$ represents the interaction effect for factors A and B at levels i and j, respectively.

Models **(7)** and **(8)** are often referred to as non-full rank models which implies that the individual parameters are non-estimable. Only functions of them are estimable, such as $\alpha_1 - \alpha_0$, $\beta_1 - \beta_0$, $(\alpha\beta)_{11} - (\alpha\beta)_{10} - (\alpha\beta)_{01} + (\alpha\beta)_{00}$ (for details see Hinkelmann and Kempthorne, 2005, Chapter 7). A more useful parameterization for **(8)**, in particular in light of our objective to analyze and evaluate interactions, is of the form

$$\begin{aligned}
f(a_0, b_0) &= \mu - \frac{1}{2}A - \frac{1}{2}B + \frac{1}{2}AB \\
f(a_1, b_0) &= \mu + \frac{1}{2}A - \frac{1}{2}B - \frac{1}{2}AB \\
f(a_0, b_1) &= \mu - \frac{1}{2}A + \frac{1}{2}B - \frac{1}{2}AB \\
f(a_1, b_1) &= \mu + \frac{1}{2}A + \frac{1}{2}B + \frac{1}{2}AB
\end{aligned} \tag{9}$$

(see Hinkelmann and Kempthorne, 2005, Chapter 7), where A and B represent the main effects of factors A and B, respectively, and AB represents the interaction between factors A and B.

In order to understand what the main effects and interaction really mean it is useful to define first *simple effects* showing the increase in response if the levels of one factor are changed and the levels of the other factor are held constant:

$$\begin{aligned} A(b_0) &= a_1b_0 - a_0b_0 \\ A(b_1) &= a_1b_1 - a_0b_1 \end{aligned} \tag{10}$$

$$\begin{aligned} B(a_0) &= a_0b_1 - a_0b_0 \\ B(a_1) &= a_1b_1 - a_1b_0 \end{aligned} \tag{11}$$

where a_ib_j in **(10)** and **(11)** represents the true effect of the treatment combination a_ib_j, i.e. ($x_1 = a_i$, $x_2 = b_j$) for i, $j = 0, 1$. Using the simple effects of factor A in **(10)** the main effect A is then defined as the average of those simple effects, i.e.

$$A = \frac{1}{2}[A(b_0) + A(b_1)] = \frac{1}{2}[a_1b_0 + a_1b_1 - a_0b_0 - a_0b_1] \tag{12}$$

Similarly, using **(11)** the main effect B is defined as

$$B = \frac{1}{2}[B(a_0) + B(a_1)] = \frac{1}{2}[a_0b_1 + a_1b_1 - a_0b_0 - a_1b_0] \tag{13}$$

To the extent that the simple effects in **(10)** and **(11)**, respectively, are not the same; i.e. to the extent that a change in one factor does depend on the level of the other factor, the difference between the simple effects provides a measure of the interaction between factors A and B, given by

$$\begin{aligned} AB &= \frac{1}{2}[A(b_1) - A(b_0)] = \frac{1}{2}[B(a_1) - B(a_0)] \\ &= \frac{1}{2}[a_1b_1 - a_1b_0 - a_0b_1 + a_0b_0] \end{aligned} \tag{14}$$

(see Hinkelmann and Kempthorne, 1994, Chapter 11). Expressions **(12)**, **(13)** and **(14)** show that the simple effects are of fundamental importance in defining main effects and interaction, and we shall return to this point in Section 7.

Concerning the interaction of the type $X \times Z$ we can proceed in a similar fashion as above. For the simplest case we assume that the function f in **(2)** is of the form $f(x, z)$. If this function can be expressed as

$$f(x, z) = f_1(x) + f_z(z) \tag{15}$$

then there is no interaction between the treatment factor x and the intrinsic factor z. This means that a change in response due to a change in x does not depend on the level of z. If interaction is present; i.e., if **(15)** does not hold, interest is only in the differential changes of the response due to change in x for each level of z. In that sense the interest in the $X \times Z$ interaction is asymmetric, whereas the interest in the $X \times X$ interaction is symmetric.

4.2 Three-factor interactions

It is instructive to extend the ideas of the previous section to three-factor interactions. For the $X \times X \times X$ interaction a possible (among many) generalization of **(6)** might be

$$f(x_1, x_2, x_3) = \alpha + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{11}x_1^2 + \beta_{12}x_1x_2 + \beta_{123}x_1x_2x_3 \tag{16}$$

It is easy to see and understand from the form of **(16)** how a change in response due to a change in x_1 now not only depends on the level of x_2, but also on the level of x_3.

Such a change is more difficult to understand if the factors are qualitative, i.e. their levels are nominal rather than quantitative, and we use the conventional overparameterized model of the form

$$f(a_i\, b_j,\, c_k) = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} \tag{17}$$

For that purpose a more useful parameterization than **(17)** may be in the form of the factorial main effects and interactions. More specifically, we shall consider the case where x_1, x_2, x_3 have two levels each, denoted by a_0, a_1 for $x_1 = A$, b_0, b_1 for $x_2 = B$ and c_0, c_1 for $x_3 = C$. Simple effects, main effects and two-factor interactions can then be defined by modifying **(10)** $-$ **(14)**. As an extension of the definition of the two-factor interaction in **(14)** it is informative to define the three-factor interaction ABC as the difference of *simple two-factor interactions* which we define as follows (see Hinkelmann and Kempthorne, 1994, Chapter 11):

$$\begin{aligned} AB(c_0) &= \frac{1}{2}\left[a_1b_1c_0 - a_1b_0c_0 - a_0b_1c_0 + a_0b_0c_0\right] \\ AB(c_1) &= \frac{1}{2}\left[a_1b_1c_1 - a_1b_0c_1 - a_0b_1c_1 + a_0b_0c_1\right] \end{aligned} \tag{18}$$

where $a_ib_jc_k$ represents the true effect of the treatment combination $(x_1 = a_i, x_2 = b_j, x_3 = c_k)$. Similar expressions can be written out for $AC(b_0)$, $AC(b_1)$, $BC(a_0)$, $BC(a_1)$. Then the three-factor interaction ABC is defined as

$$\begin{aligned} ABC &= \frac{1}{2}\left[AB(c_1) - AB(c_0)\right] \\ &= \frac{1}{2}\left[AC(b_1) - AC(b_0)\right] \\ &= \frac{1}{2}\left[BC(a_1) - BC(a_0)\right] \end{aligned} \tag{19}$$

The formulation **(19)** shows that the three-factor interaction is a measure of the extent that the interaction between any two factors depends on the level of the remaining factor. This definition is informative with respect to the evaluation of three-factor interactions, and we shall return to this point in Section 7.

Using the definitions of main effects, two-factor interactions and three-factor interaction a useful parameterization as an alternative to **(17)** can then be written as (see Hinkelmann and Kempthorne, 2005, Chapter 7):

$$f(a_i,\, b_j,\, c_k) = \mu \pm \frac{1}{2}\,A \pm \frac{1}{2}\,B \pm \frac{1}{2}\,C \pm \frac{1}{2}\,AB \pm \frac{1}{2}\,AC \pm \frac{1}{2}\,BC \pm \frac{1}{2}\,ABC \tag{20}$$

where the sign for A, B, C is "+" if the corresponding factor is present at the "high" level, e.g. a_1 for A, and "$-$" if that factor is present at the "low" level, e.g. a_0 for A, and the sign for the interactions is obtained by multiplying the signs for the corresponding main effects. As an example of **(20)** consider

$$f(a_1,\, b_0,\, c_1) = \mu + \frac{1}{2}\,A - \frac{1}{2}\,B + \frac{1}{2}\,C - \frac{1}{2}\,AB + \frac{1}{2}\,AC - \frac{1}{2}\,BC - \frac{1}{2}\,ABC$$

The other type of three-factor interaction which is of interest is of the form $X \times X \times Z$. If A and B represent the treatment factors and E the intrinsic factor, and if we consider again the special case of two levels for each factor, then

$$ABE = \frac{1}{2}\left[AB(e_1) - AB(e_0)\right] \tag{21}$$

Compared to **(19)** we have chosen in **(21)** a narrow representation of ABE. This expresses the fact that if ABE exists, interest is only in the simple AB interactions. Again, this asymmetry of interest distinguishes the $X \times X \times Z$ interaction from the $X \times X \times X$ interaction, where the interest is symmetric with respect to all three possible simple two-factor interactions.

5 NATURE OF INTERACTIONS

The reader may have noticed that we have not yet mentioned interactions of the types $X \times U$ and $X \times X \times U$. Even though these interactions can, in principle, be treated in the same way as those discussed in Section 4, they are quite different in nature. They are typically used as the error term or part of the error term in the analysis of variance, e.g. treatment $\times$ block interaction. It is for this reason that they are generally assumed to be negligible.

Unfortunately, the assumption of the absence of $X \times U$ interaction may not always be true even though it is a plausible assumption in most cases. It is obvious that if the $X \times U$ interaction term in the ANOVA table is used as the error term, then the existence of such interaction, which we might call nuisance interaction, will lead to an inflation of the error and hence to a reduction in the sensitivity (power) of subsequent tests of significance.

In the absence of pure error it is not possible to perform an overall test for $X \times U$ interaction. Only tests for special types of interaction have been developed. The most often referred to test is Tukey's one-degree-of-freedom test for nonadditivity (Tukey, 1949). A generalization of this test was developed by Mandel (1961, 1964). For a general description of these test see Hinkelmann and Kempthorne (1994, Section 9.6). We shall return to this point in Section 8 and provide another method of checking for the possible existence of $X \times U$ interaction.

One possible reason for $X \times U$ interaction to arise may be because of unequal variances of the observations. In that case one of the variance stabilizing transformations (see Hinkelmann and Kempthorne, 1994, Section 6.10) may be used to remove such heterogeneity. Often this also removes the interaction in question.

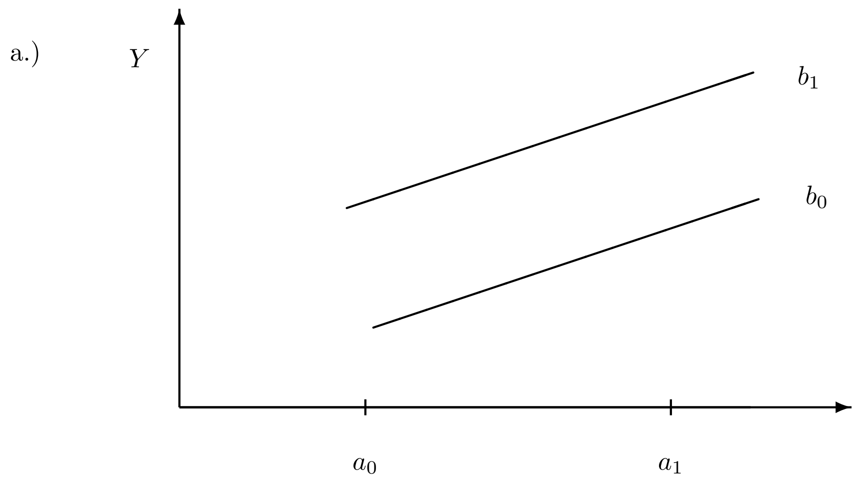
In contrast to the $X \times U$ interaction the $X \times X$ interaction is often of great importance and, in fact, desirable. For example, the application of a herbicide increased the effect of phosphorus on the seed yield of a particular variety of chick-peas (Petersen, 1994). Many experiments are performed with the specific aim to investigate whether interaction between various treatment factors exist. And if they do exist, it is important to determine the nature of such interactions.

For the special case of two factors with two levels each we give a graphical display of the different types of interaction in Figure 1.

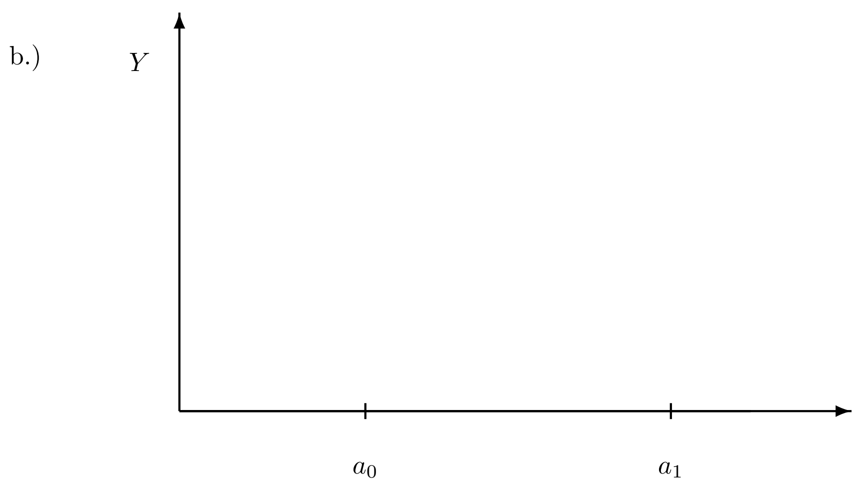
Figure 1 a.) shows that the effect of changing factor A from a_0 to a_1 on the response Y is the same regardless of the level of factor B, b_0 or b_1. This means of course, that there is no interaction. In Figure 1 b.) the corresponding change does depend on the level of B, but because the change is in the same (positive) direction Hinkelmann and Kempthorne (1994) referred to this type of interaction as codirectional interaction. The interaction depicted in Figure 1 b.) shows that the change in response is in opposite directions for the two levels of B. Hinkelmann and Kempthorne (1994) referred to this type of interaction as antidirectional interaction.

It should be clear how such interaction plots can be generalized to situations where the factors have more than two levels. The picture may not always be so clear, however, if the factors have more than two levels, but for parts of the interaction the distinction between codirectional and antidirectional will be important with respect to the interpretation and tests of main effects. Contrary to the often stated "rule" that in the presence of interaction no tests for main effects should be performed, we take the view that such tests are useful and meaningful when codirectional interaction is present, but not when antidirectional interaction is present. In the latter case it is essential to consider simple effects.

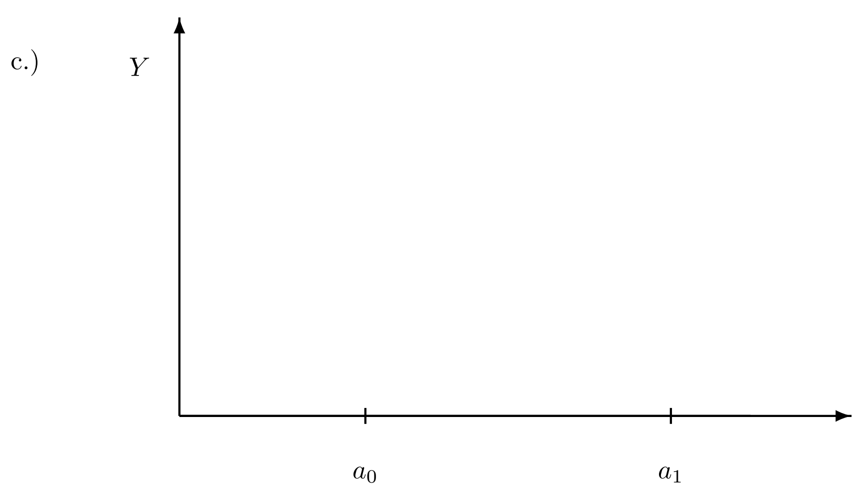
Essentially the same arguments given above for the $X \times X$ interactions apply also to the $X \times Z$ interactions. The one important difference is that the nature of the $X \times Z$ interaction is asymmetric in the following sense: The Z main effects can never be tested, even if there is no $X \times Z$ interaction or if that interaction is codirectional. The reason for this, of course, is the fact that the intrinsic factors represent (purposely introduced) blocking factors which can never be tested (Hinkelmann and Kempthorne, 1994, Chapter 9). The often expressed wish: "I want to know whether there are differences among the levels of Z" deserves usually the reply: "No, you know already that there are differences among those levels; rather, you want to know whether the levels of X behave differently for the different levels of Z". It is thus a question concerning $X \times Z$ interaction, and then the comments given above apply; i.e. for no interaction or codirectional interaction the X-main effects can be tested, for antidirectional interaction the simple effects for the X-factor at the different levels of Z need to be looked at.



a.) No interaction



b.) Codirectional interaction



c.) Antidirectional interaction

Figure 1: Types of Interaction

6 DISSECTING INTERACTIONS

The discussion in Sections 4 and 5 has already provided some indication how interactions might be analyzed and evaluated. The first step, of course, is to obtain tests for all interactions included in the underlying model for a particular experimental design. If an interaction is detected at a reasonable P-value (which might be chosen as high as .10), then we need to investigate its nature by looking at simple effects or simple interactions. We may do this by resorting to graphical and/or analytical methods.

The most basic graphical method for two-factor interaction is the interaction plot. It is simply a graphical display of the two-way table of the means (or least squares (LS) means) for the level combinations of two factors. For $X \times X$ interaction we would consider the $a_i \times a_j$ table of each pair of factors $(x_i,\, x_j)$, where x_i has a_i levels and x_j has a_j levels $(i \neq j)$. The means of x_i, connected by a line, will be displayed as a graph, separately for each level of x_j (or vice versa), representing the interaction plot. The graphs in Figure 1 represent an example where x_i and x_j have two levels each.

Similarly for $X \times Z$ interaction the interaction plot is a graphical display of the $a_i \times b_j$ table of the means for the level combinations of x_i, the treatment factor, and z_j, the intrinsic factor. The graph of the x_i-means would be displayed for every level of z_j, emphasing the asymmetry in the nature of this type of interaction.

An analytical method in conjunction with the interaction plot involves testing the equality of the x_i-means for every level of x_j or z_j in the context of the ANOVA table. In SAS this is simply accomplished by using the SLICE operator (for an illustration see Section 7). Its main function is to investigate for which levels of x_j or z_j the x_i-means (or LS means) change. It does not give any information about the direction of any possible change, but that can be obtained from the interaction plot.

One can imagine many different scenarios that may lead to $X \times X$ or $X \times Z$ interaction, and these in turn may lead to different follow-up procedures. The SLICE option mentioned above is one of them. This method may show that one or two levels of x_2 or z are mainly responsible for the $X \times X$ or $X \times Z$ interaction, respectively. In that case we may want to exclude those levels and perform the analysis without the data for those levels, leaving us, of course, with a somewhat restricted inference space.

If in particular an $X \times Z$ interaction is antidirectional or contains elements that are antidirectional, then it may be preferable to reanalyze the data separately for each level of z. This may lead to cleaner interpretation of the data, but one disadvantage is that for each of the resulting ANOVAs there are fewer d.f. for error as compared to the overall ANOVA.

It is, in general, more difficult to dissect three-factor interactions $X \times X \times X$ and $X \times X \times Z$. As mentioned above, we may want to look at simple two-factor interactions, say $x_1 \times x_2$ for each level of x_3, or $x_1 \times x_2$ for each level of z, respectively. Unfortunately, the SLICE option in SAS cannot be used for this purpose as it provides only a test for the equality of the means in the $x_1 \times x_2$ two-way table, for each level of x_3 (or z). Instead, we need to compute for each such two-way table of means (LS means) the interaction mean square, say $\mathrm{MS}(x_1 \times x_2 \mid x_3)$, and obtain the F-ratio

$$F(x_3) = \frac{\mathrm{MS}(x_1 \times x_2 \mid x_3)}{\mathrm{MS(E)}}$$

for each level of x_3 (or z), where MS(E) is the error mean square from the overall ANOVA table. For the special case where x_1 and x_2 have two levels each, we provide an explicit method in Section 7.

An alternative method is, again, to perform separate analyses, in particular if the $X \times X \times Z$ interaction is significant. The results may then lead to different actions for different levels of z.

7 ANALYZING INTERACTIONS - THE EXAMPLE REVISITED

We shall now return to the example described in Section 3 and illustrate some of the points mentioned in the preceding sections, including a complete analysis of a data set for this experiment. In order to provide the reader with the opportunity to recreate the analyses described below or to try alternative ways based on our discussion, we give the data (NOTE: these are not the original data from the actual experiment) in the form of the SAS input statement, in Table 2.

TABLE 2
DATA INPUT FOR PRUNING EXPERIMENT

```
options nodate pageno=1;
data pruning;
input V$ Block A$ B$ C Y@@;
datalines;
Am 1 F H 1 530  Am 1 F L 2 581  Am 1 M H 3 548  Am 1 M L 4 572
Am 2 F H 1 523  Am 2 F L 2 570  Am 2 M H 3 532  Am 2 M L 4 571
Am 3 F H 1 528  Am 3 F L 2 586  Am 3 M H 3 539  Am 3 M L 4 608
Am 4 F H 1 516  Am 4 F L 2 604  Am 4 M H 3 553  Am 4 M L 4 587
Am 5 F H 1 558  Am 5 F L 2 639  Am 5 M H 3 563  Am 5 M L 4 615
Am 6 F H 1 582  Am 6 F L 2 657  Am 6 M H 3 580  Am 6 M L 4 640
Ha 1 F H 1 534  Ha 1 F L 2 582  Ha 1 M H 3 554  Ha 1 M L 4 619
Ha 2 F H 1 538  Ha 2 F L 2 578  Ha 2 M H 3 543  Ha 2 M L 4 602
Ha 3 F H 1 563  Ha 3 F L 2 599  Ha 3 M H 3 567  Ha 3 M L 4 618
Ha 4 F H 1 567  Ha 4 F L 2 601  Ha 4 M H 3 601  Ha 4 M L 4 629
Ha 5 F H 1 547  Ha 5 F L 2 600  Ha 5 M H 3 607  Ha 5 M L 4 655
Ha 6 F H 1 582  Ha 6 F L 2 636  Ha 6 M H 3 602  Ha 6 M L 4 677
Co 1 F H 1 551  Co 1 F L 2 604  Co 1 M H 3 572  Co 1 M L 4 644
Co 2 F H 1 545  Co 2 F L 2 591  Co 2 M H 3 584  Co 2 M L 4 647
Co 3 F H 1 558  Co 3 F L 2 600  Co 3 M H 3 587  Co 3 M L 4 642
Co 4 F H 1 569  Co 4 F L 2 614  Co 4 M H 3 597  Co 4 M L 4 665
Co 5 F H 1 598  Co 5 F L 2 648  Co 5 M H 3 618  Co 5 M L 4 660
Co 6 F H 1 612  Co 6 F L 2 651  Co 6 M H 3 638  Co 6 M L 4 699
Fe 1 F H 1 575  Fe 1 F L 2 610  Fe 1 M H 3 590  Fe 1 M L 4 655
Fe 2 F H 1 554  Fe 2 F L 2 630  Fe 2 M H 3 605  Fe 2 M L 4 638
Fe 3 F H 1 576  Fe 3 F L 2 648  Fe 3 M H 3 608  Fe 3 M L 4 643
Fe 4 F H 1 595  Fe 4 F L 2 653  Fe 4 M H 3 631  Fe 4 M L 4 656
Fe 5 F H 1 609  Fe 5 F L 2 652  Fe 5 M H 3 641  Fe 5 M L 4 686
Fe 6 F H 1 597  Fe 6 F L 2 652  Fe 6 M H 3 660  Fe 6 M L 4 689
Pi 1 F H 1 600  Pi 1 F L 2 661  Pi 1 M H 3 625  Pi 1 M L 4 702
Pi 2 F H 1 606  Pi 2 F L 2 641  Pi 2 M H 3 635  Pi 2 M L 4 675
Pi 3 F H 1 610  Pi 3 F L 2 643  Pi 3 M H 3 642  Pi 3 M L 4 670
Pi 4 F H 1 609  Pi 4 F L 2 672  Pi 4 M H 3 653  Pi 4 M L 4 684
Pi 5 F H 1 632  Pi 5 F L 2 694  Pi 5 M H 3 669  Pi 5 M L 4 723
Pi 6 F H 1 655  Pi 6 F L 2 714  Pi 6 M H 3 676  Pi 6 M L 4 727
;
run;
```

The reader will notice that in Table 2 in addition to the pruning factors A and B (see Section 3) we have introduced a pseudo-factor C with four levels ($1 = (F, H)$, $2 = (F, L)$, $3 = (M, H)$, $4 = (M, L)$). This factor will be used later in the analysis.

The preliminary analysis follows essentially model **(3)**, except that the three-factor interaction $AB\beta$ is assumed to be negligible and thus used as the error term. The SAS input statement is given in Table 3 and the output is given in Table 4. The results in Table 4 show that certainly the $B * \mathrm{Block}(V)$ interaction is non-significant ($P = .29$) and most likely the $A * \mathrm{Block}(V)$ interaction is also negligible ($P = .16$). We therefore pool these terms with the $A * B * \mathrm{Block}(V)$ interaction, and thus the entire $X \times U$ interaction is considered to be negligible. For the subsequent analysis the $X \times U$ interaction serves as the error term with its 75 d.f.

TABLE 3
PRELIMINARY ANALYSIS
(INPUT STATEMENTS)

```
proc glm data=pruning LS=75;
class V Block A B C;
model Y=V Block(V) A B
A*B V*A V*B V*A*B A*Block(V) B*Block(V)/SS3/SS3;
run;
```

TABLE 4
PRELIMINARY ANOVA

The GLM Procedure)

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	94	261663.3833	2783.6530	30.58	<.0001
Error	25	2275.4167	91.0167		
Corrected Total	119	263938.8000			

R-Square	Coeff Var	Root MSE	Y Mean
0.991379	1.556578	9.540266	612.9000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
V	4	102743.0500	25685.7625	282.21	<.0001
Block(V)	25	50019.2500	2000.7700	21.98	<.0001
A	1	18451.2000	18451.2000	202.72	<.0001
B	1	78540.8333	78540.8333	862.93	<.0001
A*B	1	108.3000	108.3000	1.19	0.2858
V*A	4	3750.7167	937.6792	10.30	<.0001
V*B	4	306.5833	76.6458	0.84	0.5117
V*A*B	4	1494.7833	373.6958	4.11	0.0108
Block*A(V)	25	3415.5833	136.6233	1.50	0.1582
Block*B(V)	25	2833.0833	113.3233	1.25	0.2939

Based on the results of Table 4 we are then led to the model

$$\begin{aligned} Y_{ijk\ell} &= \mu + V_i + \beta_{ij} + A_k + B_\ell + (AB)_{k\ell} \\ &\quad + (VA)_{ik} + (VB)_{i\ell} + (VAB)_{ik\ell} \\ &\quad + e_{ijk\ell} \end{aligned} \tag{22}$$

The SAS input statements for the ANOVA using model **(22)** and for some follow-up procedures are given in Table 5. Among these are the slice options 'LSMEANS $A*B/$ SLICE $= B$ SLICE $= A$' and 'LSMEANS $A*V$ $B*V$ $A*B*V/$ SLICE $= V$'. With regard to the $A*B$ interaction, the slice option tests whether the simple effects for A and B, respectively, are significant. In general, the slice option tests the equality of the LS means for one factor at the different levels of the other factor. With regard to the $V*A$ and $V*B$ interactions, the option 'SLICE $= V$' affects that we test whether the simple effects of A and B are significant for each level of V (considering the asymmetry of the $X \times Z$ interaction, these two are the only slice options of interest here). We note that we did include the option 'LSMEANS $A*B*V/$ SLICE $= V$' only to show that this would result to test whether the four LS means for (F, H), (F, L), (M, H), and (M, L) are different from each other for every level of V, and that is of no interest to us.

TABLE 5
ANOVA FOR MODEL (22)
(INPUT STATEMENTS)

```
proc glm data=pruning;
class V Block A B;
model Y=V Block(V) A B A*B A*V B*V A*B*V/SS3;
lsmeans A B A*B/slice=B slice=A;
estimate 'Main effect A' A -1 1;
estimate 'Main effect B' B -1 1;
lsmeans  A*V B*V A*B*V/slice=V;
run;
```

We now turn to the analysis as presented in Table 6 and make the following comments:

1. The P-values for V and Block (V) should be ignored, since under randomization theory no significance tests for block effects, i.e., effects of intrinsic and non-specific factors, are permissible (see Hinkelmann and Kempthorne, 1994, Chapter 9).

2. The $A*B$ interaction is non-significant ($P = .33$). Thus there is no real need to invoke the slice option. We have included it, however for purposes of illustration. Using the $A*B$ LS means and **(10)** and **(11)** we find the estimates of the simple effects to be

$$\begin{aligned}\widehat{A}(H) &= 600.67 - 573.97 = 26.7\\ \widehat{A}(L) &= 649.93 - 627.03 = 22.9\end{aligned}$$

   and

$$\begin{aligned}\widehat{B}(F) &= 627.03 - 573.97 = 53.06\\ \widehat{B}(M) &= 649.93 - 600.67 = 49.26\end{aligned}$$

   and all are significantly different from zero ($P < .0001$), and so are the estimates of the main effects $\widehat{A} = 24.8$ and $\widehat{B} = 51.17$ with standard error 1.95. At the same time, the test for $A*B$ interaction indicates that the simple effects for A as well as those for B are not different from each other, leading to near-parallel lines in the interaction plot.

3. Among the $X \times Z$ interactions, $A*V$ and $A*B*V$ are significant ($P < .0001$ and $P = .015$, respectively). The results for the $A*V$ interaction sliced by V indicate that only the simple A-effects for variety Am are not different from each other, whereas the estimates of the simple A-effects for the other four varieties are of the same order of magnitude, around 30, as can be seen from the $A*V$ LS means. This interaction is clearly a codirectional interaction, and hence considering the overall

TABLE 6
ANOVA FOR MODEL (22)
WITH FOLLOW-UP TESTS

The GLM Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	44	255414.7167	5804.8799	51.07	<.0001
Error	75	8524.0833	113.6544		
Corrected Total	119	263938.8000			

R-Square	Coeff Var	Root MSE	Y Mean
0.967704	1.739417	10.66088	612.9000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
V	4	102743.0500	25685.7625	226.00	<.0001
Block(V)	25	50019.2500	2000.7700	17.60	<.0001
A	1	18451.2000	18451.2000	162.34	<.0001
B	1	78540.8333	78540.8333	691.05	<.0001
A*B	1	108.3000	108.3000	0.95	0.3321
V*A	4	3750.7167	937.6792	8.25	<.0001
V*B	4	306.5833	76.6458	0.67	0.6118
V*A*B	4	1494.7833	373.6958	3.29	0.0154

Least Squares Means

A	Y LSMEAN
F	600.500000
M	625.300000

B	Y LSMEAN
H	587.316667
L	638.483333

A	B	Y LSMEAN
F	H	573.966667
F	L	627.033333
M	H	600.666667
M	L	649.933333

A*B Effect Sliced by B for Y

B	DF	Sum of Squares	Mean Square	F Value	Pr > F
H	1	10693	10693	94.09	<.0001
L	1	7866.150000	7866.150000	69.21	<.0001

A*B Effect Sliced by A for Y

A	DF	Sum of Squares	Mean Square	F Value	Pr > F
F	1	42241	42241	371.66	<.0001
M	1	36408	36408	320.34	<.0001

Least Squares Means

V	A	Y LSMEAN
Am	F	572.833333
Am	M	575.666667
Co	F	595.083333
Co	M	629.416667
Fe	F	612.583333
Fe	M	641.833333
Ha	F	577.250000
Ha	M	606.166667
Pi	F	644.750000
Pi	M	673.416667

V*A Effect Sliced by V for Y

V	DF	Sum of Squares	Mean Square	F Value	Pr > F
Am	1	48.166667	48.166667	0.42	0.5170
Co	1	7072.666667	7072.666667	62.23	<.0001
Fe	1	5133.375000	5133.375000	45.17	<.0001
Ha	1	5017.041667	5017.041667	44.14	<.0001
Pi	1	4930.666667	4930.666667	43.38	<.0001

V	B	Y LSMEAN
Am	H	546.000000
Am	L	602.500000
Co	H	585.750000
Co	L	638.750000
Fe	H	603.416667
Fe	L	651.000000
Ha	H	567.083333
Ha	L	616.333333
Pi	H	634.333333
Pi	L	683.833333

V*B Effect Sliced by V for Y

V	DF	Sum of Squares	Mean Square	F Value	Pr > F
Am	1	19154	19154	168.52	<.0001
Co	1	16854	16854	148.29	<.0001
Fe	1	13585	13585	119.53	<.0001
Ha	1	14553	14553	128.05	<.0001
Pi	1	14702	14702	129.35	<.0001

V	A	B	Y LSMEAN
Am	F	H	539.500000
Am	F	L	606.166667
Am	M	H	552.500000
Am	M	L	598.833333
Co	F	H	572.166667
Co	F	L	618.000000
Co	M	H	599.333333
Co	M	L	659.500000
Fe	F	H	584.333333
Fe	F	L	640.833333
Fe	M	H	622.500000
Fe	M	L	661.166667
Ha	F	H	555.166667
Ha	F	L	599.333333
Ha	M	H	579.000000
Ha	M	L	633.333333
Pi	F	H	618.666667
Pi	F	L	670.833333
Pi	M	H	650.000000
Pi	M	L	696.833333

V*A*B Effect Sliced by V for Y

V	DF	Sum of Squares	Mean Square	F Value	Pr > F
Am	3	19822	6607.277778	58.13	<.0001
Co	3	24235	8078.277778	71.08	<.0001
Fe	3	19195	6398.486111	56.30	<.0001
Ha	3	19725	6575.152778	57.85	<.0001
Pi	3	19675	6558.277778	57.70	<.0001

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Main effect A	24.8000000	1.94640219	12.74	<.0001
Main effect B	51.1666667	1.94640219	26.29	<.0001

A main effect is appropriate. Furthermore, the slice operation shows also that the $A \times V$ interaction comes about only because of the different behavior of variety Am. If we were to drop Am from the analysis, there would be no $A \times V$ interaction. Since the $B \times V$ interaction is not significant, the $B * V$ slice operation is not really needed. The $B * V$ LS means show that the estimators of the simple B-effects are all about the same order of magnitude, around 50.

4. Concerning the $A \times B \times V$ interaction, we have included the slice operation with $A*B*V$ effect sliced by V only to demonstrate that this operator does not produce the desired results for three-factor interactions. We would like to compare the simple $A \times B$ interactions, but the results of the slice operation indicate that this procedure tests, for each variety, the equality to the four $A*B$ LS means, as indicated by $DF = 3$. To achieve our objective we now use the factor C introduced earlier, noting that the contrast vector $(1, -1, -1, 1)$ for C describes the $A \times B$ interaction. More specifically, we use the SAS input statement as given in Table 7 and the results in Table 8, as follows.

TABLE 7
INVESTIGATING A*B*V INTERACTION
(INPUT STATEMENTS)

```
proc glm data=pruning;
class V Block C;
 model Y=V Block(V) C V*C;
estimate 'A*B for Am' C 1 -1 -1 1 V*C 1 -1 -1 1/divisor=2;
estimate 'A*B for Co' C 1 -1 -1 1 V*C 0 0 0 0 1 -1 -1 1/divisor=2;
estimate 'A*B for Fe' C 1 -1 -1 1 V*C 0 0 0 0 0 0 0 0 1 -1 -1 1/divisor=2;
estimate 'A*B for Ha' C 1 -1 -1 1 V*C 0 0 0 0 0 0 0 0 0 0 0 0 1 -1 -1 1/divisor=2;
estimate 'A*B for Pi' C 1 -1 -1 1 V*C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 -1 -1
1/divisor=2;
run;
```

5. We note that from Tables 6 and 8 it follows that the following relationships among sums of squares exist:

$$SS(C) = SS(A) + SS(B) + SS(A * B)$$

and

$$SS(V * C) = SS(V * A) + SS(V * B) + SS(V * A * B)$$

with 3 and 12 d.f., respectively. From $SS(C) + SS(V * C)$ with 15 d.f. the estimate statements in Table 7 isolate 5 d.f. which specify the simple $A * B$ interactions for each variety. The results in Table 8 indicate that only the $A * B$ interactions for Am is clearly significant ($P = .022$). The $A * B$ interaction for Fe is borderline significant ($P = .044$), whereas the other $A * B$ interactions are not significant. Informally, this point is shown graphically in the half-normal plot of Figure 2, where the absolute values of the interaction effects are plotted, with those for Fe and Am deviating from the straight line (for an explanation of this technique see Section 8). This shows again that the variety Am behaves somewhat differently than the other varieties. A closer look at the $V * A * B$ LS means confirms this finding as the highest yield for all varieties except Am is achieved for the (M, L) treatment combination. For Am the highest yield is obtained for the (F, L) treatment combination, but the difference between the yields for (M, L) and (F, L) is relatively small.

6. Thus, the overall conclusion from this study shows over a wide range of pear varieties the method of light pruning with many leaders will lead to the best results. Possible exceptions are for varieties similar to Am, where light pruning with few leaders may produce slightly better results.

TABLE 8
COMPARING SIMPLE A*B INTERACTIONS

The GLM Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	44	255414.7167	5804.8799	51.07	<.0001
Error	75	8524.0833	113.6544		
Corrected Total	119	263938.8000			

R-Square	Coeff Var	Root MSE	Y Mean
0.967704	1.739417	10.66088	612.9000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
V	4	102743.0500	25685.7625	226.00	<.0001
Block(V)	25	50019.2500	2000.7700	17.60	<.0001
C	3	97100.3333	32366.7778	284.78	<.0001
V*C	12	5552.0833	462.6736	4.07	<.0001

Parameter	Estimate	Standard Error	t Value	Pr > |t|
A*B for Am	-10.1666667	4.35228761	-2.34	0.0222
A*B for Co	7.1666667	4.35228761	1.65	0.1038
A*B for Fe	-8.9166667	4.35228761	-2.05	0.0440
A*B for Ha	5.0833333	4.35228761	1.17	0.2465
A*B for Pi	-2.6666667	4.35228761	-0.61	0.541

The follow-up procedures as described above are done within the context of the overall analysis using model **(22)**. And this is the method we recommend in general. One reason for proceeding in this way is that all inferences are based on the same error term, namely MS(Error) from the overall ANOVA, usually based on a sufficient number of degrees of freedom.

An alternative procedure, however, might be to perform separate analyses for each level of one intrinsic factor or for each level combination of several intrinsic factors. Since in our example the $V \times A \times B$ interaction is significant, we might be led to five analyses, each based on the model

$$Y_{jk\ell}^{(i)} = \mu^{(i)} + \beta_j^{(i)} + A_k^{(i)} + B_\ell^{(i)} + (AB)_{k\ell}^{(i)} + e_{jk\ell}^{(i)} \tag{23}$$

for $i = 1, 2, \ldots, 5$, with input statements given in Table 9. Although we would be able to make then recommendations separately for each variety, it becomes more difficult to arrive at statistically sound overall conclusions. We shall not provide all the details of the five analyses here, but summarize in Table 10 the most important information.

TABLE 9
SEPARATE ANALYSES
(INPUT STATEMENTS)

```
proc sort data=pruning;
by V;
run;

proc glm data=pruning;
by V;
class Block A B;
model Y= Block A|B;
lsmeans A B A*B/slice=B;
run;
```

TABLE 10: INFORMATION FROM SEPARATE ANALYSES

Level of V	MS(Error) 15 d.f.	P-Values A	B	$A*B$
Am	132.48	.556	.0001	.047
Co	60.54	.0001	.0001	.039
Fe	130.19	.0001	.0001	.075
Ha	145.15	.0001	.0001	.318
Pi	99.91	.0001	.0001	.523

Based on the results presented in Table 10 we make the following comments:

$1'$. The MS(Error) appear to be quite different for the different varieties. This might be used as an argument that separate analyses are really called for. However, testing the equality of variances, using Hartley's F-max test with $F_{max} = 145.15/60.54 = 2.4$ (Hartley, 1950) leads to the result that the MS(Error) are not significantly different at the 5% level. Hence, from this point of view, the overall analysis is justified.

$2'$. The main effects for A are significant except for variety Am. This agrees with the finding in 3. above.

$3'$. All main effects for B are significant.

$4'$. The $A * B$ interactions for varieties Am, Co, and Fe are significant or borderline significant, but are not significant for varieties Ha and Pi. The only real deviation from our earlier findings in 5. above is the result for Co. This is due to the fact that MS(Error) for Co is quite small. Referring to the LS means for $V * A * B$ in Table 6, we see that all significant interactions are codirectional.

$5'$. Overall, the conclusions from the separate analyses would have been the same as from the overall analysis (see 6. above).

The conclusions from the statistical discussion in this section can be summarized as follows:

$1''$. Start the analysis with as large a model as possible, i.e., make as few as possible assumptions about possibly negligible interactions.

$2''$. Investigate first all $X \times U$ interactions to establish the "correct" error term.

$3^{''}$. Analyze all $X \times X$, $X \times Z$, and $X \times X \times Z$ interactions by means of graphical and analytical methods. The graphical methods are often most useful for purposes of illustration and interpretation, but the results must be confirmed by analytical methods, usually in the form of follow-up tests to the ANOVA.

$4^{''}$. The question whether to use one overall analysis or separate analyses does not always have an easy answer. We advocate the former to the extent possible. This analysis will in general be more powerful, but may require more sophisticated approaches to analyzing interactions. This complication may be avoided by using separate analyses. The possible disadvantage then are lack of power and conclusions that may lack a certain amount of generality. Thus, this question may have to be decided on a case by case basis.

8 BLOCKING BY INTRINSIC FACTOR ONLY

The experiment described in Section 3 and analyzed in Section 7 uses a randomized complete block design with a nested block structure, i.e., blocks nested within varieties or, more generally, a non-specific factor nested within an intrinsic factor. In this section we shall consider the situation if the experiment had consisted of just one block for each variety. In that case variety is the only blocking factor. In other words, the intrinsic factor is the only blocking factor.

The typical approach to analyzing data from such an experiment would be to assume that the treatment $\times$ block interaction is negligible and then to use the model

$$Y_{ijk} = \mu + V_i + A_j + B_k + (AB)_{jk} + e_{ijk} \tag{24}$$

We know, however, from the analysis of the larger experiment that the $V \times A$ and $V \times A \times B$ interactions were significant there. Thus, the assumptions that lead to model **(24)** may not be appropriate.

In general, we do not have this kind of insight, but whenever an intrinsic factor is used as a blocking factor careful consideration must be given to possible existence of $X \times Z$ interactions. Usually such considerations have to be based on subject matter knowledge rather than on statistical arguments since there may not exist a test for the $X \times Z$ interaction. Tests proposed by e.g. Tukey (1949) and Mandel (1961), mentioned in Section 5, are possible statistical approaches to this dilemma.

In our example, the treatments have a factorial structure. Therefore the $X \times Z$ interaction can be divided into $A \times V$, $B \times V$, and $A \times B \times V$ interactions. This provides us with a choice whether to assume that all three interaction components or only one or two of them are negligible.

We shall propose here a different approach to this problem. We shall adapt the method of half-normal plots which was proposed by Daniel (1959) to identify non-zero effects in a saturated fraction of a 2^n factorial. Saturated in this context means that the design does not provide any d.f. for error. This is the same situation here if we are not willing, a priori, to assume that some of the $X \times Z$ interactions are negligible. In order to use this method (for a description see Daniel, 1959; Zahn, 1975; Hinkelmann and Kempthorne, 2005, Section 13.9) for our purpose we need to partition the $X \times Z$ interaction into single-d.f.-contrasts. In general, if X has ν_x d.f. and Z has ν_z d.f., then $X \times Z$ has $\nu_x \cdot \nu_z$ d.f. Thus, there will be $\nu_x \cdot \nu_z$ contrasts which will have to be orthonormal , i.e. orthogonal and normalized, for this procedure to work. We shall use our example to describe how this can be accomplished. The general procedure should then become obvious.

We have $\nu_x = 3$ and $\nu_z = 4$, where the 3 d.f. for X are represented by those for A, B, and AB, and the 4 d.f. for Z by four comparisons among the five varieties, denoted by $V1$, $V2$, $V3$, $V4$ say. For $V1$, $V2$, $V3$, $V4$ we choose the complete set of four orthogonal polynomials among the five varieties. The contrast coefficients for these orthogonal polynomials are given in Table 11 (see Hinkelmann and Kempthorne, 1994, Section 7.4). We should note that these contrasts have no particular meaning here since the levels of V are nominal, but that they were chosen conveniently for mathematical purposes only; other contrasts could have been chosen just as well as long as they are orthogonal. The seven sets of contrast coefficients are given in Table 11, labelled $V1$, $V2$, $V3$, $V4$, A, B, AB. The coefficients for the 12 contrasts belonging to the $X \times Z$ interactions are then simply obtained by multiplying the coefficients for the corresponding X and Z contrasts. For example, the coefficients for the contrast $V1A$ is obtained by multiplying elementwise the coefficients for $V1$ and A. For the set of contrasts labelled $V1A$, $V2A$, …, $V4AB$ we also give the normalizing divisor (ND), which is the square root of the sum of the squared coefficients. We then obtain the contrast estimates and plot their absolute values on half-normal probability paper. The results are given in Figure 3.

Inspection of Figure 3 shows that the absolute values for $V1A$, $V2A$, $V3A$, $V4A$, and $V2B$ do no lie on the line going through the smaller contrast values. This implies that, at least informally, these contrasts are not negligible and, hence, probably should not be included in the error term. In other words, rather than using

$$SS(X \times Z) = SS(A \times V) + SS(B \times V) + SS(AB \times V)$$

as the error sum of squares, as suggested by model **(24)**, we might want to use

$$SS(\text{Error}) = SS(B \times V) + SS(AB \times V) - SS(V2B) \tag{25}$$

since

$$SS(A \times V) = \sum_{i=1}^{4} SS(ViA)\,,$$

and $SS(V2B) = (\widehat{V2B})^2$ (see Hinkelmann and Kempthorne, 1994, Section 7.23). However, since the contrast $V2B$ represents only one d.f. of the 4 d.f. interaction $B \times V$, it may be quite appropriate from a practical point of view to use, instead of **(25)**, the final error term

$$SS(\text{Error}^*) = SS(B \times V) + SS(AB \times V)\,. \tag{26}$$

The form of the error term **(26)** implies that the data should be analyzed according to the model

$$Y_{ijk} = \mu + V_i + A_j + B_k + (AB)_{jk} + (VA)_{ij} + e^*_{ijk} \tag{27}$$

The analysis of the data using model **(27)** is presented in Table 12. We comment briefly on their results:

1. The main effects A and B are significant with $P < .0001$.
2. The interaction $A * B$ is not significant ($P = .2023$).
3. The $A*V$ interaction is significant ($P = .0031$) as suggested already by the half-normal plot of Figure 3.
4. Slicing the $A * V$ interaction indicates that the simple $A-$effects for variety Am are not significantly different from each other. Inspection of the $V * A$ LS means shows that the $A \times V$ interaction is codirectional.

9 CONCLUSIONS AND RECOMMENDATIONS

In many experimental and observational studies the question how to deal with interactions among various factors is an important one. The easiest way out is to suppress the interactions by assuming − tacitly or expressly − that they are negligible. This may, of course, also be a very dangerous way to proceed as such assumptions may be wrong. In fact, we have argued that interactions involving treatment factors or treatment and intrinsic factors are often important and need at least be investigated. To this end we have presented and discussed some analytical and graphical procedures. The interpretation of existing interactions then rests on understanding the meaning of interaction, not only in a statistical sense but also in the context of the experimental situation. This leads us to the following recommendations:

1. Design the experiment in accordance with the underlying question for the experiment (see Hinkelmann and Kempthorne, 1994, Chapter 2).
2. Provide a list of all the statistical and subject-matter factors important in the experiment and identify treatment, intrinsic, and non-specific factors.

TABLE 11: SINGLE - D.F. CONTRAST PARTITION

Y	558	639	563	615	547	600	607	655	598	648	618	660	609	652	641	686	632	694	669	723	
V	Am				Ha				Co				Fe				Pi				ND
$A*B$	FH	FL	MH	ML	FH	FL	MH	ML	FH	FL	MH	ML	FH	FL	MH	ML	FH	FL	MH	ML	
$V1$	−2	−2	−2	−2	−1	−1	−1	−1	0	0	0	0	1	1	1	1	2	2	2	2	
$V2$	2	2	2	2	−1	−1	−1	−1	−2	−2	−2	−2	−1	−1	−1	−1	2	2	2	2	
$V3$	−1	−1	−1	−1	2	2	2	2	0	0	0	0	−2	−2	−2	−2	1	1	1	1	
$V4$	1	1	1	1	−4	−4	−4	−4	6	6	6	6	−4	−4	−4	−4	1	1	1	1	
A	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	
B	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	
AB	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	
$V1A$	−2	−2	2	2	−1	−1	1	1	0	0	0	0	1	1	−1	−1	2	2	−2	−2	6.325
$V2A$	2	2	−2	−2	−1	−1	1	1	−2	−2	2	2	−1	−1	1	1	2	2	−2	−2	7.483
$V3A$	−1	−1	1	1	2	2	−2	−2	0	0	0	0	−2	−2	2	2	1	1	−1	−1	6.325
$V4A$	1	1	−1	−1	−4	−4	4	4	6	6	−6	−6	−4	−4	4	4	1	1	−1	−1	16.733
$V1B$	−2	2	−2	2	−1	1	−1	1	0	0	0	0	1	−1	1	-1	2	−2	2	-2	6.325
$V2B$	2	−2	2	−2	−1	1	−1	1	−2	2	−2	2	−1	1	−1	1	2	−2	2	−2	7.483
$V3B$	−1	1	−1	1	2	−2	2	−2	0	0	0	0	−2	2	−2	2	1	−1	1	−1	6.325
$V4B$	1	−1	1	−1	−4	4	−4	4	6	−6	6	−6	−4	4	−4	4	1	−1	1	−1	16.733
$V1AB$	−2	2	2	−2	−1	1	1	−1	0	0	0	0	1	−1	−1	1	2	−2	−2	2	6.325
$V2AB$	2	−2	−2	2	−1	1	1	−1	−2	2	2	−2	−1	1	1	−1	2	−2	−2	2	7.483
$V3AB$	−1	1	1	−1	2	−2	−2	2	0	0	0	0	−2	2	2	−2	1	−1	−1	1	6.325
$V4AB$	1	−1	−1	1	−4	4	4	−4	6	−6	−6	6	−4	4	4	−4	1	−1	−1	1	16.733

3. Write out as complete a linear model as possible (accounting for all d.f. for the complete data set) by including all possible interactions.

4. Using subject-matter knowledge, decide which interactions can be assumed to be negligible.

5. Investigate whether statistical tests exist for the remaining interactions, or how they can be performed for important interactions, for example by using a generalized randomized block design instead of a randomized complete block design (see Hinkelmann and Kempthorne, 1994, Chapter 9).

6. Analyze and interpret the interactions using some of the techniques discussed above.

7. Remember that a statistical analysis can only be as good as the underlying experimental design.

TABLE 12
ANOVA FOR MODEL (27)
WITH FOLLOW-UP PROCEDURES

The GLM Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	39278.40000	3570.76364	59.79	<.0001
Error	8	477.80000	59.72500		
Corrected Total	19	39756.20000			

R-Square	Coeff Var	Root MSE	Y Mean
0.987982	1.225336	7.728195	630.7000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
V	4	19287.70000	4821.92500	80.74	<.0001
A	1	3380.00000	3380.00000	56.59	<.0001
B	1	14045.00000	14045.00000	235.16	<.0001
A*B	1	115.20000	115.20000	1.93	0.2023
V*A	4	2450.50000	612.62500	10.26	0.0031

Least Squares Means

A	Y LSMEAN	Standard Error	Pr > |t|
F	617.700000	2.443870	<.0001
M	643.700000	2.443870	<.0001

B	Y LSMEAN	Standard Error	Pr > |t|
H	604.200000	2.443870	<.0001
L	657.200000	2.443870	<.0001

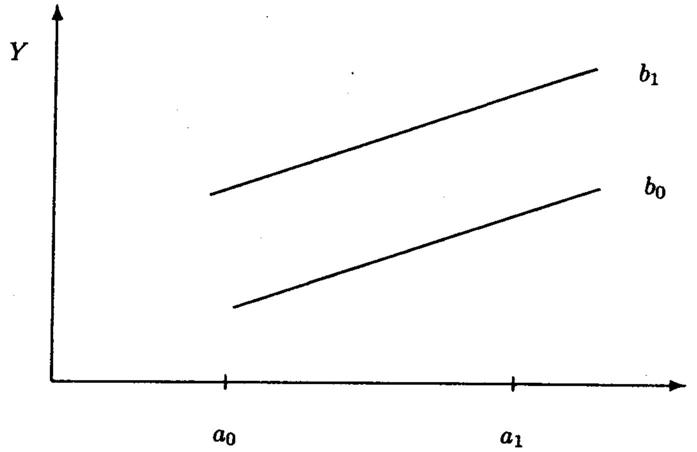
A	B	Y LSMEAN	Standard Error	Pr > |t|
F	H	588.800000	3.456154	<.0001
F	L	646.600000	3.456154	<.0001
M	H	619.600000	3.456154	<.0001
M	L	667.800000	3.456154	<.0001

V	A	Y LSMEAN	Standard Error	Pr > |t|
Am	F	598.500000	5.464659	<.0001
Am	M	589.000000	5.464659	<.0001
Co	F	623.000000	5.464659	<.0001
Co	M	639.000000	5.464659	<.0001
Fe	F	630.500000	5.464659	<.0001
Fe	M	663.500000	5.464659	<.0001
Ha	F	573.500000	5.464659	<.0001
Ha	M	631.000000	5.464659	<.0001
Pi	F	663.000000	5.464659	<.0001
Pi	M	696.000000	5.464659	<.0001

V*A Effect Sliced by V for Y

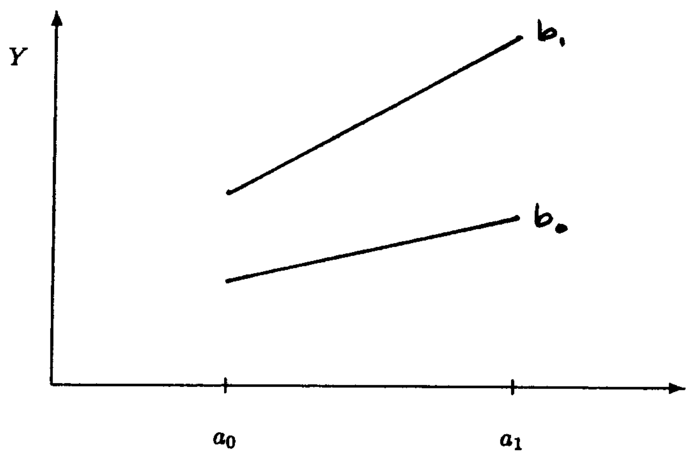
V	DF	Sum of Squares	Mean Square	F Value	Pr > F
Am	1	90.250000	90.250000	1.51	0.2539
Co	1	256.000000	256.000000	4.29	0.0722
Fe	1	1089.000000	1089.000000	18.23	0.0027
Ha	1	3306.250000	3306.250000	55.36	<.0001
Pi	1	1089.000000	1089.000000	18.23	0.0027

a.)



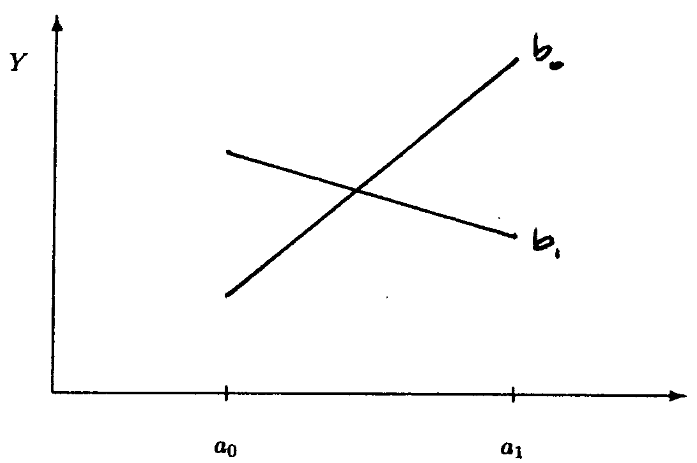
a.) No interaction

b.)



b.) Codirectional interaction

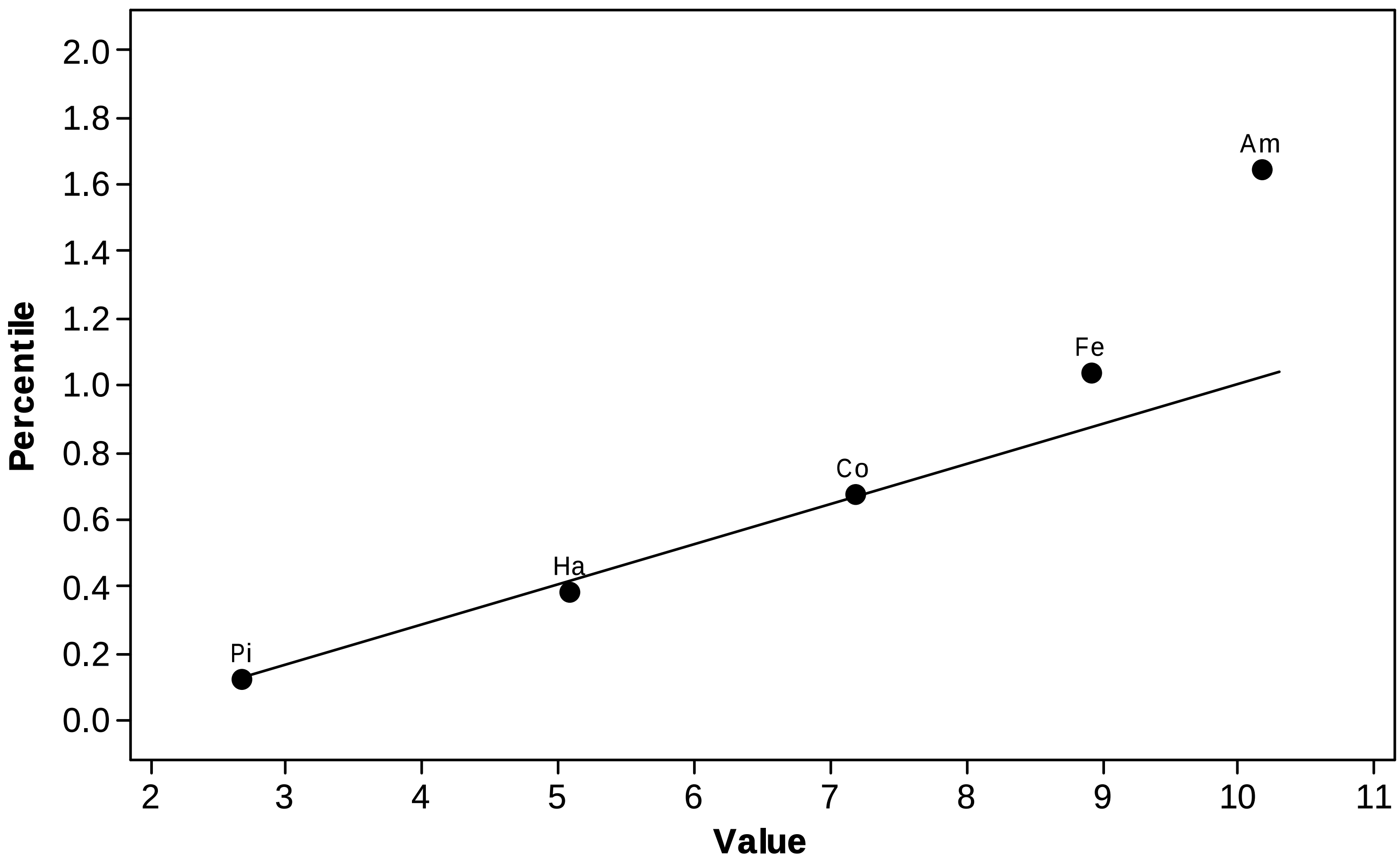
c.)



c.) Antidirectional interaction

Figure 1: Types of Interaction

Figure 2: Half-Normal Plot of AB Interactions



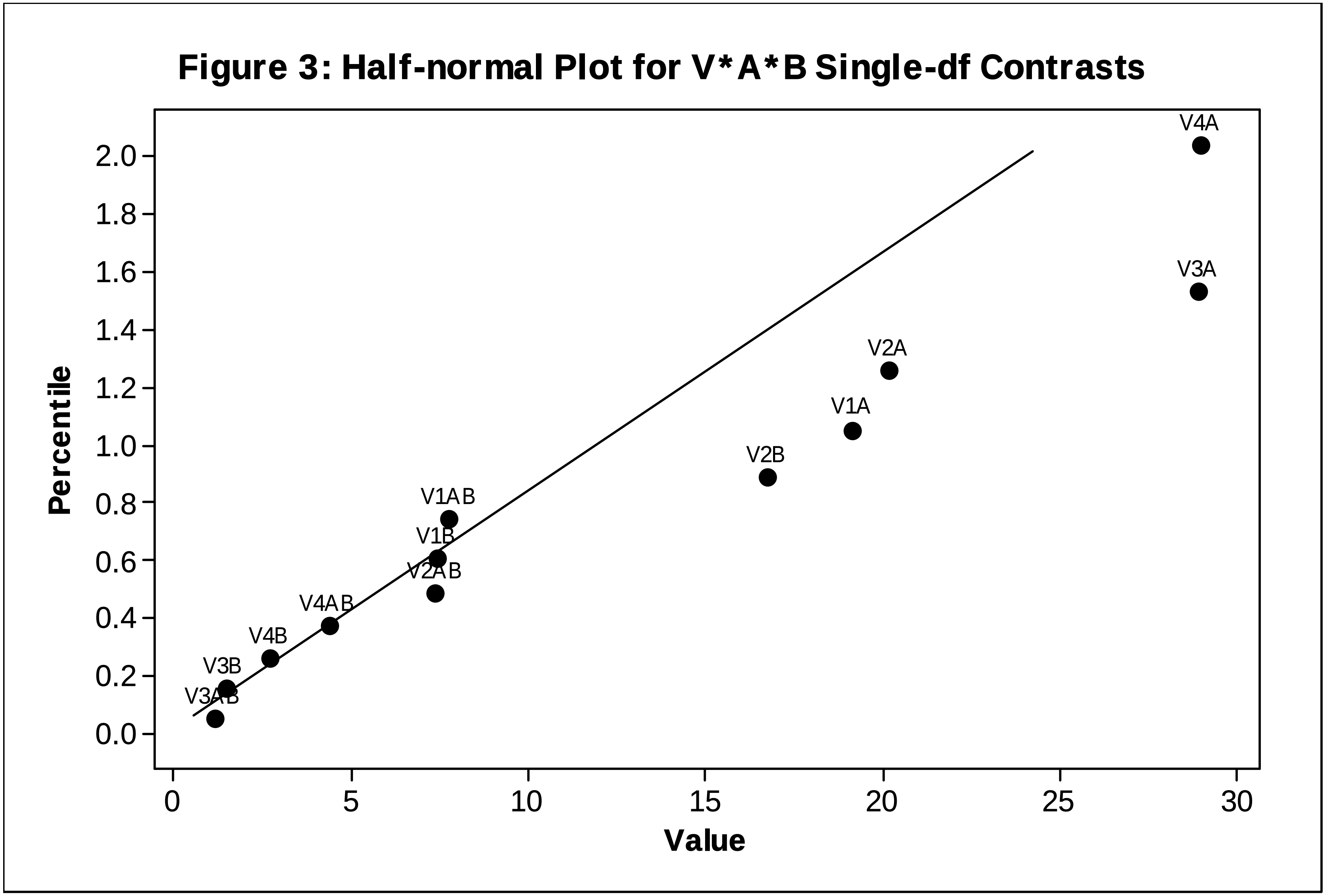


Figure 3: Half-normal Plot for V*A*B Single-df Contrasts
Percentile
Value
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
V4A
V3A
V2A
V1A
V2B
V1AB
V1B
V2AB
V4AB
V4B
V3B
V3AB

References

[1] Cox, D. R. (1984). Interaction (with discussion). *Intern. Statist. Rev.*, **52**, 1-32.

[2] Daniel, C. (1959). Use of half-normal plots in interpreting factorial two-level expriments. *Technometrics*, 1**1**, 311-341.

[3] Hartley, H. O. (1950). The maximum F-ratio as a short-cut test for heterogeneity of variance. *Biometrika* **37**, 308-312.

[4] Hinkelmann, K. and O. Kempthorne (1994). ***Design and Analysis of Experiments. Vol. 1***. New York: Wiley.

[5] Hinkelmann K. and O. Kempthorne (2005). ***Design and Analysis of Experiments. Vol. 2***. New York: Wiley.

[6] Mandel, J. (1961). Non-additivity in two-way analysis of variance. *J. Amer. Statist. Assoc.*, **56**, 878-888.

[7] Mandel, J. (1964). ***The Statistical Analysis of Experimental Data***. New York: Dover Publications.

[8] Pearce, S. C. (1953). Field experimentation with fruit trees and other perennial plants. *Commonwealth Agric. Bus. Techn. Communc.*, **23**.

[9] Pearce, S. C. (1983). ***The Agricultural Field Experiment***. Chichester: Wiley.

[10] Petersen, R. G. (1994). ***Agriculture Field Esperiments: Design and Analysis***. New York: Dekker.

[11] SAS Institute (2000). *SAS® Proprietary Software*, Release 8.2. Cary, N.C.: SAS Institute.

[12] Tukey, J. W. (1949). One degree of freedom for nonadditivity. *Biometrics*, **5**, 232-242.

[13] Zahn, D. A. (1975). Modifications of the revised critical values for the half-normal plot. *Technometrics*, **17**, 189-200.