

Answers to Selected Exercises

CHAPTER 1

1.1 (a) $z = \frac{-13.5}{4.5\sqrt{\frac{2}{20}}} = -9.49$ The p -value is essentially zero.

(b) $z = \frac{-13.5}{4.5\sqrt{\frac{1}{30} + \frac{1}{10}}} = -8.21$ The p -value is still essentially zero.

The numbers show the importance of having a state of statistical control when experimentation is performed.

1.2 (b)

1.3 In order to maximize the amount of information that results from a designed experiment, we need to know as much about the experimental region as possible and how the response values will vary over that region. We gain that information from a well-designed first experiment. Thus, we could have almost certainly done a better job of designing the first experiment if we had that knowledge. Generally more than one experiment will be performed however, so not having good information about the experimental region before performing the first experiment would be a handicap only when experimentation is extremely expensive.

1.4 No, there is an obvious pairing of the data, as would occur in a taste-testing experiment, for example.

1.6 Repeated readings occur when readings are successively made at a fixed experimental condition, with no resetting of factor levels and no randomness involved. With *replication*, we can speak of replicating an entire experiment (in this chapter) or replicating certain factor-level combinations (in subsequent

chapters). With replication there is a random ordering of experimental runs and resetting factor levels.

- 1.7 Because there is no randomness associated with the assignment of the last treatment to the last experimental unit.
- 1.8 With at most four observations per level, there would be no way to check the assumptions of normality and equal variances.
- 1.12 The world is nonlinear and experiments with two levels, which are appropriate for linearity, won't always give satisfactory results.
- 1.14 Such information generally does not exist. There would be no point in experimentation only if all other important factors were at their optimal levels and were held constant and the factor of interest were then varied "from one level to another level." Such a state of affairs will very rarely exist.
- 1.15 This suggests that a process was probably out of control and affected the values of the response variable. My suggestion would be to check the critical processes, and if a process problem is detected, fix the problem if possible and rerun the experiment.
- 1.21 The imbalance does not create a problem as this can be analyzed as a one-factor ANOVA problem with unequal sample sizes. There is not a significant operator effect because the value of the F -statistic is 1.76 and the associated p -value is .210.
- 1.23 An increase in the variance for either or both levels will make it harder to detect a difference between the two means because the estimate of the (assumed) common variance will be inflated.

CHAPTER 2

- 2.1 $\hat{A}_i = \bar{y}_i - \bar{\bar{y}}$ so $\sum_{i=1}^k \hat{A}_i = \sum_{i=1}^k (\bar{y}_i - \bar{\bar{y}}) = \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij} / n - k \bar{\bar{y}}$
 $= \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij} / n - k \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij} / nk = 0$
- 2.4 (a) There were 31 observations and 4 levels of the factor; 31/4 is not an integer.
 (b) 8, 8, 8, and 7 for levels 1–4, respectively.
- 2.6 The necessary assumptions are the same (and it was stated that the factor is fixed, so that ANOM can be used). There is no evidence that the population variances would differ (by any significant amount) and the normality assumption could not be tested with such a small amount of data. The averages are 16.4, 17, 16.2, and 20.4. The results differ because three of the averages do not differ greatly, with the last average considerably greater than the other three. Although multiple comparisons are usually not performed unless the null

hypothesis of equal population means is rejected, in this case there is reasonably strong evidence that there is a difference involving the last mean.

2.7 $df(\text{error}) = 4 + 6 + 5 = 15.$

2.9 (a) $SS(\text{treatments}) = 40$, so $MS(\text{treatments}) = 2$. Since $F = 4$, $MS(\text{error}) = 0.5$. Then $SS(\text{error}) = 0.5(12) = 6$ and $SS(\text{total})$ thus equals 46.

(b) Since $F_{2,12} = 4$ has a p -value of .047, the hypothesis would (barely) be rejected using $\alpha = .05$.

2.10 6.67

2.13 (a) The smallest possible value is zero; that will occur when the total of the observations for the two levels is the same. The largest possible value is infinity.

(b) Level One: 2, 4, 5, 6, 9, and 11
Level Two: 1, 3, 6, 7, 10, and 10

CHAPTER 3

3.1 Statement (b) suggests that a design with blocks should have been used but we would not expect the treatment totals to be about the same. The motive for the experimentation with these particular treatments would have to be examined.

3.3 The five factors would be randomly assigned to the experimental units for each block.

3.4 Consider the 5×5 Latin square:

A	B	C	D	E
B	A	D	E	C
C	E	A	B	D
D	C	E	A	B
E	D	B	C	A

If we delete the last two rows, the design is not balanced because, for example, the (A, B) pair will occur in the first and second columns, but most other pairs will occur only once. Hence, the design is not balanced, and is therefore not a balanced incomplete block design and thus not a Youden design.

3.8 (a)

A	B	C	D	E
B	A	D	E	C
C	E	A	B	D
D	C	E	A	B
E	D	B	C	A

(b) $Y_{ijk} = \mu + A_i + R_j + C_k + \epsilon_{ijk} \quad \begin{matrix} i = 1, 2, \dots, 5 & j = 1, 2, \dots, 5 \\ k = 1, 2, \dots, 5 \end{matrix}$

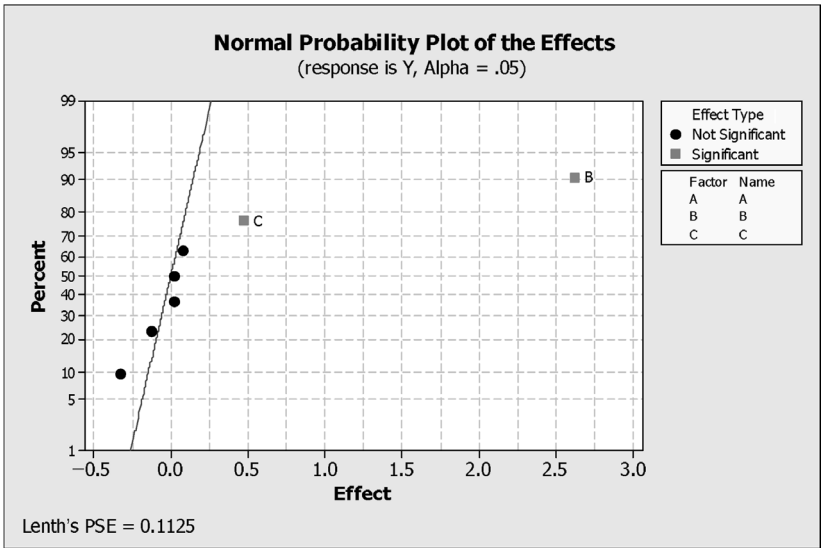
The model assumes no interactions between rows and columns nor between treatments and either rows or columns. These assumptions should be checked with appropriate graphs, such as shown in Figure 3.2. The error term is assumed to have a normal distribution, but there is no way to test that with only one observation per cell.

- 3.10** It then becomes a two-factor problem with blocking, but it was not designed as such. That is, an appropriate randomization scheme was not used, so a Latin square analysis should still be performed, with the row or column factor (whichever it is) viewed as a factor of interest.
- 3.16** The numbers suggest that an RCB design should have probably been used instead of a Latin square design since Columns is not significant.
- 3.20** (a) $F_{4,12} = 4$ gives a p -value of .027.
 (b) Neither rows nor columns are significant, so a design with blocking seems unnecessary.
- 3.21** 16.565

CHAPTER 4

- 4.1** The main effect of A is estimated by $\frac{1}{2}(a + ab - (1) - b)$. The conditional effects of A are $ab - b$ and $a - (1)$, so $ab - b = (1) - a$ since the conditional effects differ only in sign. Then $ab - b + a - (1) = 0$, so the estimate of the A effect is zero.
- 4.5** The conditional effects will differ very little since the conditional effects are equal to the main effect plus and minus the interaction effect.
- 4.6** There is no multicollinearity in a 2^3 design.
- 4.8** The problem with that line of thinking is that the main effect coefficients in the model will not well represent the conditional effects when there are large interactions.
- 4.11** A useful order would be $A, B, AB, C, AC, BC, D, AD, BD, CD, ABC, ABD, ACD, BCD, ABCD$. The trick would be to construct the display in such a way that it wouldn't look cluttered.
- 4.12** (a) The B effect estimate is 0.5 but this should not be reported to management because the clear interaction effect causes the conditional effects to be -2 and 3 . It would be better to report those numbers, with the caveat that each number is computed as the difference of just two numbers.
 (b) 4

- 4.14 $(250 - 312.5)/62.5 = -1$ and $(375 - 312.5)/62.5 = 1$
- 4.15 The first is a design for three factors, each at two levels; the second is a design for two factors, each at three levels.
- 4.18 Anything of higher order than a four-factor interaction almost certainly won't be real, so it would be safe to use such interactions in estimating sigma. In general, however, it is best to estimate sigma with replications, not with interactions.
- 4.31 The normal probability plot (see below) suggests that the *B* and *C* main effects are the significant effects.



- 4.32 Homogeneity of variance must also be assumed with ANOM because sigma is estimated the same way with ANOM as it is estimated with ANOVA.
- 4.33 Yes, ANOM is a graphical *t*-test when there are two levels for each factor. The degrees of freedom is 32, assuming that sigma is estimated from the replicates.

CHAPTER 5

- 5.4 The treatment combination that is incorrectly grouped with the others is *acd*. This has an odd number of letters in common with the interaction in the defining relation, whereas the other four treatment combinations (which would be used

with the other 12 treatment combinations) have an even number of letters in common with the treatment combination.

- 5.5** If $I = ABDE$, 2 two-factor interactions would be confounded in pairs and the design would thus be resolution IV. A resolution V design can be constructed by using the defining relation $I = ABCDE$, which is how the design should be constructed.
- 5.8** (a) The defining relation is $I = ABC$. This result could be obtained using the approach given by Bisgaard (1993) or simply by trial and error for a small design such as this.
 (b) Yes, this is obviously a 2^{4-1} design, which should be constructed using $I = ABCD$, which would confound main effects with three-factor interactions.
 (c) With the design that is given, $A = BC$, $B = AC$, and $C = AB$.
- 5.10** There would be 32 design points because this would have to be a 2^{6-1} design since E is aliased with only one effect.
- 5.13** The conditional effects of factor D will differ only slightly because AB is confounded with CD , so when the data are split on factor C , the conditional effects would have to be close, by definition.
- 5.15** A minimum aberration design won't necessarily be the design that gives the maximum number of clear effects.
- 5.18** This is a 2^{5-2} design. It was constructed in an unorthodox manner since the generator for factor E is $E = -BC$ and the generator for D was $-ABC$. Letting $E = BC$ would be customary and this would reverse all of the signs in the fifth column of the design. (Alternatively, the generator $E = AC$ could have been used.) It would also be customary to use $D = AB$. If this were used, the first four signs in the fourth column would be the same and the last four would be the reverse of what they are in the problem.
- 5.19** Highly fractionated designs are usually low-resolution designs, for which there is a risk of one or more factors being declared significant because they are confounded with significant interactions.
- 5.24** The third column is the column that would be used to generate the fifth factor if the design were constructed as a 2^{5-1} design. We can see that something is awry, however, because the bottom half of the column is the same as the top half of the column. That isn't what happens when a 2^{5-1} design is constructed, as the bottom half should be the mirror image of the top half. Therefore, each of the eight numbers in the top half of the third column should be changed to the other number in order for the design to be resolution V. As it now

stands, the design is resolution IV as it can be shown (such as by creating fictitious response values and analyzing the data) that the defining relation is $I = ABCE$.

- 5.25** We can see by inspection that $E = BCD$, so $I = BCDE$ is the defining relation, which is the same result that would be obtained using the methodology of Section 5.6. Of course this is not the best way to construct the design, however, as using $I = ABCDE$ will maximize the resolution of the design.
- 5.33** No, a 2^{9-1} design would have 256 design points, which would be extremely wasteful. Estimating the nine main effects and 36 two-factor interactions requires only 46 design points. Any design that has a much larger number of design points is wasteful.
- 5.40** It is possible to construct a 2_{III}^{9-5} design that has four words of length three in the defining relation. The design used by Hsieh and Goodwin (1986) has considerably more words of that length. Thus, the latter is not the minimum aberration design. A superior design would have the following generators: $E = ABC$, $F = ABD$, $G = ACD$, $H = BCD$, and $J = ABCD$.

CHAPTER 6

- 6.3** With $I = ABC$, the full alias structure is

$$A = BC = AB^2C^2$$

$$B = AB^2C = AC$$

$$C = ABC^2 = AB$$

- 6.4** The configuration should be the same, which should be expected because of the X configuration. Furthermore, the middle-level values are the same (and are equal) for each factor, and the set of response values for the high level of each factor is the same, as is the set of response values for the low level of each factor. Thus, the factor designations are interchangeable relative to the interaction plot.
- 6.11** The designs will be different. For $D = ABC^2$ we have $x_1 + x_2 + 2x_3 - x_4 = 0$ so $x_1 + x_2 + 2x_3 + 2x_4 = 0$ and $I = ABC^2D^2$. For $D = A^2B^2C$ we have $2x_1 + 2x_2 + x_3 + 2x_4 = 0$. Adding $3x_3$ to each side of the equation, and then dividing by 2, we obtain $x_1 + x_2 + 2x_3 + x_4 = 0$ so $I = ABC^2D$. Thus, the designs are different because the defining relations are different. There may not be any reason to prefer one design over the other, unless there are certain combinations of factor levels that cannot be used together for physical reasons.

- 6.13** No, the fractions are different. This can be verified by constructing the fraction for which the sum is zero (mod 3) for each design and seeing that the column for factor D is different in each fraction.
- 6.15** The full second-order model would have 35 terms (7 linear, 7 quadratic, 21 interaction terms), which would require 98 degrees of freedom. Even a 3^{7-3} design is only resolution IV and that has 81 points. A 3^{7-2} design would have 243 points and generally be quite impractical, except in computer experiments when runs are inexpensive.

CHAPTER 7

7.3 We would compute the sum of squares for $D(H)$ as

$$\begin{aligned} 3 \sum_{i=1}^2 \sum_{j=1}^2 (\overline{H_i D_j} - \overline{H_i})^2 &= 3[(16/3 - 35/6)^2 + (19/3 - 35/6)^2 \\ &\quad + (23/3 - 50/6)^2 + (27/3 - 50/6)^2] \\ &= 4.167 \end{aligned}$$

as given in the table.

7.8 The full degrees of freedom breakdown is:

Source	df
A	4
B (within A)	5
C (within B)	5
D (Error)	5
Total	19

7.13	Source	df
	Schools	1
	Classroom (School)	6

CHAPTER 8

- 8.5** The interaction plot suggests that the high level of the factor is better than the low level, but the line for the high level is not as close to being horizontal as we would prefer.
- 8.6** Although slopes close to zero are desirable, if both slopes are close to zero there is no clear advantage of one level over the other one. This is still preferable,

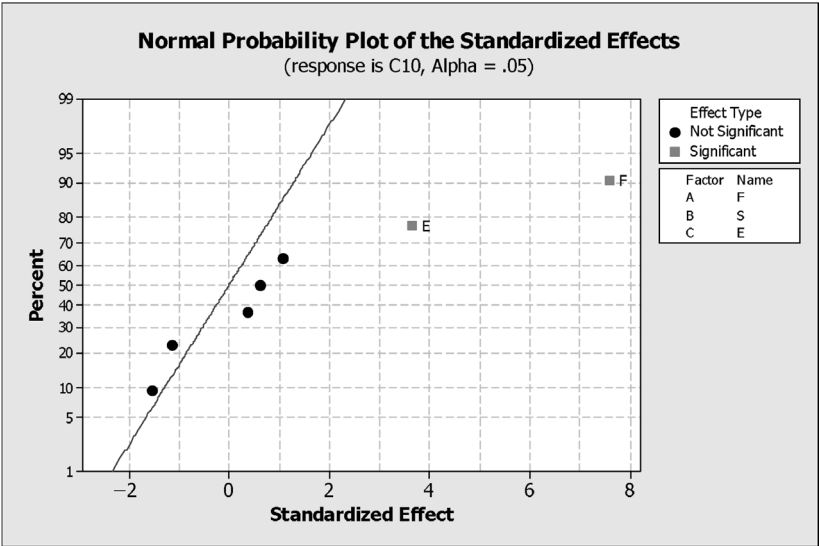
however, to an interaction profile with the slope of one line close to +1 and the slope of the other line close to −1 as this would tell us that neither level of the control factor would be a good choice, relative to the particular noise factor that is plotted on the horizontal axis.

- 8.7 Since noise factors cannot be controlled during production, we want to find levels of control factors for which the response is virtually the same regardless of the levels of the uncontrollable noise factors.
- 8.9 The design in Exercise 8.8 does not have the same weakness because it is not a 2^{k-p} design. Since the design has 12 points and is a saturated design, this is also a Plackett–Burman design (see Section 13.4.1) and thus has the same complex alias structure as a Plackett–Burman design. Those designs are resolution III designs. The L_{12} design is suitable for factor screening but not for investigating control \times noise interactions or any other type of interaction.

CHAPTER 9

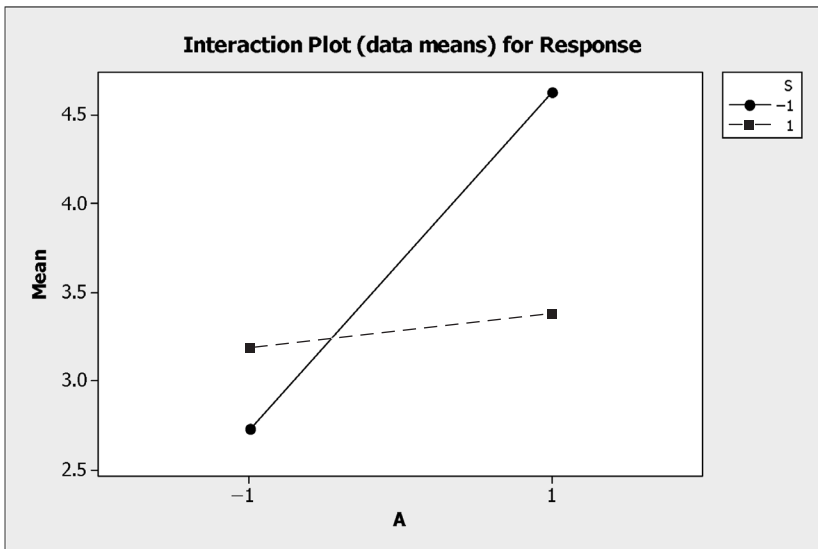
- 9.5 Since there are five factors each at two levels and 32 design points, this is analogous to a 2^5 design. Thus, there will be no error terms so two normal probability plots will have to be constructed. The data can be initially analyzed as if a 2^5 design had been used for the purpose of obtaining the effect estimates.

The normal plot for the whole plot effects below shows that the *E* (egg powder) and *F* (flour) effects are significant.



Analyzing the subplot effects is a chore, however, because it requires fitting a nonhierarchical model and producing the normal plot, which most software programs will not allow. We can produce a normal probability plot of effect estimates, however, by using the effect estimates from the 2^5 analysis and then testing manually the points that are well off the line. The two points that stand out in that plot are the A main effect and the SA interaction. The normal plot of Box and Jones (2000) showed these effects to be significant. Manual computations using Lenth's approach (Section 4.10) produce results that are in agreement with those results, with the B effect close to being borderline.

The magnitude of the SA interaction is evidenced by the following graph.

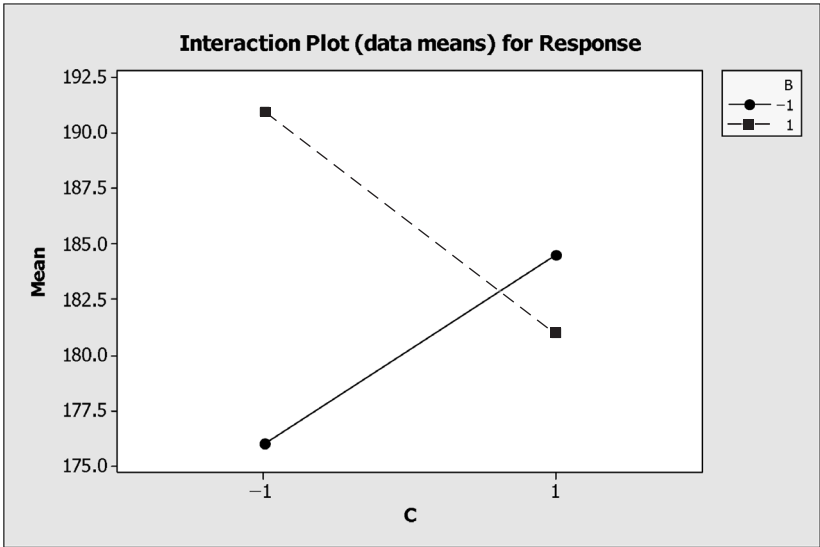


The graph also shows that the high level of factor S is clearly the best from a robust product standpoint because the variability in the response over the two levels of the noise factor A is clearly much less than when S is at the low level.

- 9.13** In order to have a split-plot design, there must be at least one whole plot treatment in addition to subplot treatments. Here there is only one treatment with two levels, selfing and outcrossing. (This observation was also made by a prominent statistician who responded on the Internet to the experimental setting that was presented. That person felt that the main concern should be a suitable model for the count data. Whether or not that should be a major concern should probably depend on the magnitude and variability of the counts, however.)
- 9.14** Zero. There would be neither enough observations nor enough factors to permit a normal probability plot estimate.

CHAPTER 10

- 10.10** The surface plot is a function of the data on the response variable, which is a random variable. If a second experiment were conducted and another surface plot displayed, using the same model form as used the first time, the surface plot will look at least slightly different from the first one.
- 10.16** The *BC* interaction plot is given below.



- We can see from the plot that the *C* effect estimate will be close to zero since the average value at high *C* is obviously close to the average value at low *C*. The conditional effects of *C*, however, are sizable. Therefore, this should not be overlooked and accordingly the (linear) effect of *C* should not be overlooked (as the experimenters have apparently done) in terms of modeling.
- 10.19** The path of steepest ascent is determined by expressing the amount of change in each factor as a multiple of a unit change in one selected factor, with the multipliers computed from the coefficients in the fitted first-order model. This cannot be done if there is an interaction term in the model because the contribution to change in the response then results from joint changes in the two or more factors that comprise the interaction term.
- 10.25** For the full CCD, all the pairwise correlations between the quadratic effect columns are $-.124$; for the half CCD the correlations are all $-.067$. Thus, the half CCD has a slightly better correlation structure, in addition to being a more economical design.

CHAPTER 11

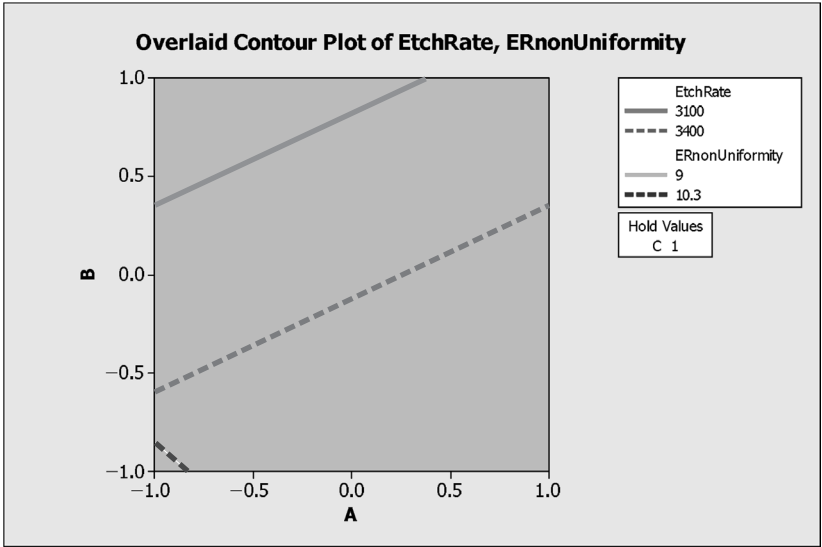
- 11.1** No, this Latin square design would be totally unsuitable when there are carryover effects because each treatment is always preceded by the same treatment when it is not used first in a treatment period (A always precedes B , D always precedes A , and so on.)
- 11.7** A 3×3 Latin square is not a good design to use when carryover effects may be present. In particular, it is not possible to construct a single Latin square of this size that will be balanced for carryover effects. As discussed in Section 3.3.5, multiple Latin squares are highly desirable, but there are restrictions on the number of Latin squares that can be used and still have balance relative to carryover effects. For example, whereas the following two Latin squares are balanced in regard to possible carryover effects

A B C	A C B
B C A	B A C
C A B	C B A

because each treatment is preceded by the same treatment in each period in the first square when the treatment is not used first in a treatment period, and by the other treatment in the second square, the balance would obviously be lost if a third square were used. More specifically, B , for example, obviously has to be preceded by A or C . If we pick one or the other and have B preceded by one of the two in the third Latin square when it is not the first treatment, then the set of three Latin squares will be unbalanced relative to the treatment that precedes B . It is easy to show that we cannot construct the third Latin square in such a way that B is preceded by A in one row and by C in another row, in addition to of course being first in the other row. Therefore, a set of three Latin squares cannot be balanced in regard to possible carryover effects.

CHAPTER 12

- 12.1** We don't know what desirability function was used because there is no discussion of it. One potential problem with the initial experiment is that some two-factor interactions were stated as being significant, which might be masking important factors that were discarded and not used in the subsequent experimentation.
- 12.8** There was only a small feasible region when factor C was set at the low level, as was shown in Figure 12.2. When C is set at the high level, there is no feasible region, as shown by the absence of a white (common) area in the contour plot given below.



CHAPTER 13

- 13.2 A split-plot design would be a logical choice, with the hard-to-change factor being the whole plot factor.
- 13.6 It is certainly true that the response values determine whether or not points are influential, but it would be unwise to create potential problems by using a design that has some high leverage values.
- 13.12 A 2^{31-26} design, which of course is resolution III.
- 13.15 The objective is to dealias effects and create a resolution IV design in the process.

CHAPTER 14

- 14.1 One possibility would be to use a trend-free design. This would protect against an out-of-control state that could be represented by a linear trend, as occurs with tool wear, for example.
- 14.7 Supersaturated designs can be useful but it is important not to use a supersaturated design that has more than small correlations between the columns of the design.

- 14.9** Even when data are inexpensive, 160 observations to estimate (at most) 15 effects can have deleterious effects because the hypothesis tests that are performed in testing for effect significance will have too much power in the sense that effects for which the estimates are quite small could be declared real effects.
- 14.12** Such information generally does not exist. It is hard enough to know which main effects can be expected to be real; pretending to know which interactions are likely to be real is apt to cause erroneous conclusions to be drawn.
- 14.15** There are many useful response surface designs that are not orthogonal designs, such as the Hoke designs, with the larger Hoke designs having small correlations. Nonorthogonality is not a major problem as long as the departure from nonorthogonality is not great.
- 14.16** Because the design runs are expensive, an OFAT design of resolution V should be strongly considered, despite the person's opposition to OFAT designs. I would attempt to be persuasive by using illustrative examples.