Designing and Running Super-Efficient Experiments: Optimum Blocking With One Hard-to-Change Factor

FRANK T. ANBARI

The George Washington University, Washington, DC 20052

JAMES M. LUCAS

J. M. Lucas and Associates, 5120 New Kent Road, Wilmington, DE 19808

This paper discusses how to run 2^k experiments for process improvement when there is one hard-to-change factor. The paper studies the different ways of running these experiments and gives practical recommendations. It shows how to block designs to get small prediction variance and low cost. It presents an algorithm to allow the selection of efficient blocking relations, in 2^k designs, where there is one hard-to-change factor and tabulates the results for 2^3 to 2^7 designs, in various block sizes. It presents methods for calculating the prediction variance and G-efficiency when there are hard-to-change factors. The calculations are demonstrated by applying them to 2^k designs, and results are tabulated for various block sizes. We show that optimally blocked split-plot designs dominate randomized designs. A blocked split-plot design is both less expensive to run, because it requires fewer resets of the hard-to-change factor, and more precise, as it gives a lower variance of prediction than a completely randomized design.

Key Words: Block Size; Cost Function; Expensive-to-Change Factors; G-Efficiency; Prediction Variance.

DESIGNED experiments are an important component of process-improvement and quality-enhancement projects. There are many practical situations where one or more factors in the experiment are hard, i.e., difficult, expensive, or time consuming to change. Oven temperature, line set-up, and formula change are typically hard-to-change factors. In practice, it is often observed that when successive runs have the same level of the hard- or expensive-to-change factor, that factor is not independently reset for each run. Therefore, a completely randomized design (CRD) is not obtained (Ju and Lucas (2002), Ganju and Lucas (1997, 1999, 2005), and Webb et al. (2004)). To get a CRD, the design must be run in a random order and each factor must be reset on each run. Running the design in a random order does not give a CRD. The design is a blocked design with the blocks determined at random and there are two or more error terms. Running the design in orthogonal (or near orthogonal) blocks gives a better design. This paper shows that more blocked designs and fewer CRDs should be run because the blocked design will be more cost efficient and more precise than a CRD.

Ganju and Lucas studied the performance of standard statistical tests when there is a hard-to-change factor. Their 1997 paper showed that the tests on easy-to-change factors are correct or nearly correct over all randomizations. However, the tests on hard-to-change factors can be far off. Their 1999 paper showed that, when the number of parameters in the model is not a small fraction of the design points, then split plotting is difficult to detect after the fact.

Dr. Anbari is an Assistant Professor in the Department of Decision Sciences. He is a Senior Member of the ASQ and an ASQ Certified Six Sigma Black Belt. His e-mail address is anbarif@gwu.edu.

Dr. Lucas is the Principal at J. M. Lucas and Associates. He is a Fellow of the ASQ and the ASA. His e-mail address is jamesm.lucas@verizon.net.

Their 2005 paper showed that the correctness of tests for easy-to-change factors over all run orders is a small consolation because, for many run orders, the tests can be far off.

Anbari (1993, 2004) and Anbari and Lucas (1994) showed that, when there is a hard-to-change factor, the design should be blocked on the hard-to-change factor. They showed that one can get designs with efficiencies greater than 100% relative to a CRD and that these designs will also be more cost effective than designs run in a random order. Designs having efficiencies greater than 100% relative to a CRD are called "super-efficient" designs.

This paper shows that 2^k experiments should be designed, run, and analyzed differently from the current practice of running the designs in a random order. We determine the prediction variance and calculate the efficiency for various ways of running the experiment. We show how proper blocking achieves efficiencies greater than 100% relative to a CRD. We discuss the cost of running the experiment and show practical block patterns that have higher G-efficiencies and higher cost efficiencies than designs run in a random order.

Some implications of this paper that can be considered a general set of guidelines in design of experiments are:

- The experimenter should pay close attention to the way the experiment will actually be run, to achieve maximum efficiency.
- Algorithmic (computer-generated) designs can give the answer to the wrong design questions.
- A Taguchi crossed array is sometimes the most cost-effective way to run an experiment (but definitely not always the way).
- Journal editors should require more details about how experiments are conducted.
- It is appropriate to conduct more split-plot experiments and fewer randomized experiments.

Hard-to-Change Factors

An important principle in the design of experiments is randomization, which guards against bias due to trends or cycles in the experiment. In a CRD, all factors that require resetting are reset for each experimental run.

However, in many designed experiments, randomization can be expensive or unfeasible because some factors in the experiment are hard to change. Examples of the presence of a hard-to-change factor can be noted in changing the temperature of the oven used to bake components (Wortham and Smith (1959)), changing the setup of the system for measuring the distance of the astronomical unit (Youden (1972)), changing the temperature of the drying furnace in the production of laminate boards (Taguchi (1984, 1987)), changing the temperature of a mold in a chemical industrial experiment (Lucas and Ju (1992), Webb et al. (2004)), and changing the powder coating formula (O'Neill and DeBrunce (2001)). In such cases, the level of the hard-to-change factors may not be reset independently whenever two or more successive runs occur with the same level of those factors (Ju and Lucas (2002)). An experiment run in a random order, but with one or more factors not reset for each run, will be referred to as a "random run order design" (Ju (1992), Ganju and Lucas (2005)) or a "randomized not reset (RNR) experiment" (Webb et al. (2004)). If the analysis is made assuming that complete randomization had occurred when in fact it did not, the analysis can be erroneous (Ganju and Lucas (1997), Webb et al. (2004)). Goos et al. (2006) discussed response surface experiments when some of the factors are not reset independently and pointed out that the result is a split-plot experimental design, and the observations in the experiment are, in many cases, correlated.

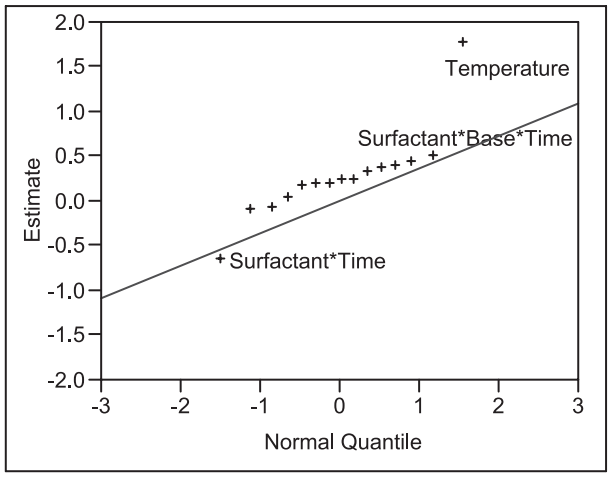
Motivating Examples

We give two examples of blocking with split-plot experiments when there is one hard-to-change factor (factor A). Example 1, 2^4 main effects and two-factor interactions model: Ju and Lucas (2002) discussed this design and showed that a 4-block design using the defining relationship I = A = BCD = ABCD is super efficient (also see Table 4). Here we describe an experimental program that used this design. In the production of a man-made fiber, a finish is usually added to the fiber at the last phase of the production process. The finish makes it easier to weave the fabric. After the weaving, it is often desirable to remove the finish to enhance the performance of the woven fabric. Our experimental product is a tightly woven industrial fabric that can be used in protective clothing. Residual finish can lubricate a projectile that is fired at the wearer of the protective clothing so the finish must be removed. Because the solvent emissions from the current dry-cleaning process were environmentally unacceptable, an alternative finish-removal process was desired. We describe

TABLE 1. Finish Removal Results (JMP 6)
Data List 2^4 Imbedded Design

Run order	Block	Temp.	Surf.	Base	Time	Finish
1	1	−1	−1	−1	1	11.7
2	1	−1	−1	1	−1	11.4
3	1	−1	1	1	1	11.4
4	1	−1	1	−1	−1	12.6
5	2	1	−1	−1	−1	14.4
6	2	1	−1	1	1	14.7
7	2	1	1	1	−1	15.5
8	2	1	1	−1	1	12.7
9	3	1	−1	−1	1	16.5
10	3	1	−1	1	−1	13.8
11	3	1	1	−1	−1	16.6
12	3	1	1	1	1	17.6
13	4	−1	−1	1	1	12.3
14	4	−1	−1	−1	−1	11.1
15	4	−1	1	1	−1	12.8
16	4	−1	1	−1	1	9.9

the lab-scale experimental program that used three experiments to demonstrate the feasibility of removing the finish by washing the fabric. The first experiment used a 2^4 experiment in 4 blocks plus additional experimental points. The essential results of the experiment are seen in the 2^4 imbedded design that we show as Table 1. The four factors are *surfactant* type, an additive (*base*) to reduce the acidity of the wash, wash *time*, and wash *temperature* (the hard-to-change factor). For proprietary reasons, the levels are coded and a linear transformation is applied to the response (residual *finish*) that was measured on 3″ by 10″ swaths of fabric. Figure 1 and Table 2 show the results of a JMP 6 analysis. The half-normal plot Figure 1 shows temperature away from the straight line and also shows the sideways "Z" shape that is typical of split-plot experiments. The temperature regression coefficient is the largest effect but it is in the "wrong" direction with more finish remaining with the hotter wash. The Table 2 mixed-model analysis for a main effects plus two-factor interactions model clearly shows two error terms but no significant factor effects (at the 0.05 level). The temperature effect ($p = 0.0589$) is not quite significant. However, in all runs, the residual *finish* is an order of magnitude higher than the level achieved by dry cleaning, so the feasibility of using washing to replace dry cleaning is not demonstrated. A second experiment that changed some factors and levels



The line is Lenth's PSE, from the estimates population

FIGURE 1. Normal Probability Plot.

gave similar results; the residual *finish* was still too high. The third experiment that washed the fibers from the swaths (after the swaths were "unwoven") was successful in reducing the residual *finish* to a low level. This completed the lab-scale program and demonstrated that washing can remove the finish. The lab-scale program also identified problems that must be addressed before washing is a feasible commercial process. Because of the tightly woven fabric, the commercial wash must use a strong force for the *surfactant* and the rinse to enable them to remove the finish.

Example 2, 2^4 main effects model: We show that an excellent blocking procedure is to use 8 blocks with defining generator I = A = BD = CD, so BC, ABC, ABD, ACD are also confounded with blocks. The 8-block experiment has fewer changes of the hard-to-change factor than a CRD (8 versus 16), and we also show (in Table 4) that it has a smaller maximum variance of prediction (and higher G-efficiency) over the experimental region than a CRD. Because the G-efficiency of the CRD is 100%, the blocked design is "super efficient" relative to the CRD. This experiment also has a higher G-efficiency and fewer expected changes than an RNR experiment that uses a random run order but does not reset the hard-to-change factor when successive runs have the same level.

Design Efficiencies

Kiefer (1959, 1961, and other papers) and others proposed the use of several optimality criteria

TABLE 2. Finish Removal Analysis Results (JMP 6 output)
Response Finish

Summary of fit

R square	0.961455
R square adj	0.884366
Root mean square error	0.913327
Mean of response	13.4375
Observations (or sum wgts)	16

Parameter estimates

Term	Estimate	Std error	DFDen	t ratio	Prob $>\|t\|$
Intercept	13.4375	0.45432	2	29.58	0.0011
Temperature	1.7875	0.45432	2	3.93	0.0589
Surfactant	0.2	0.228332	3	0.88	0.4456
Base	0.25	0.228332	3	1.09	0.3536
Time	−0.0875	0.228332	3	−0.38	0.7271
Temperature ∗ surfactant	0.175	0.228332	3	0.77	0.4992
Temperature ∗ base	−0.075	0.228332	3	−0.33	0.7641
Temperature ∗ time	0.2375	0.228332	3	1.04	0.3747
Surfactant ∗ base	0.4375	0.228332	3	1.92	0.1512
Surfactant ∗ time	−0.65	0.228332	3	−2.85	0.0653
Base ∗ time	0.4	0.228332	3	1.75	0.1781

REML variance component estimates

Random effect	Var ratio	Var component	Std error	95% lower	95% upper	Pct of total
Block	0.7397602	0.6170833	0.8430004	−1.035198	2.2693642	42.521
Residual		0.8341667	0.6810942	0.2676928	11.596639	57.479
Total		1.45125				100.000

−2 log likelihood = 46.533254622

to compare experimental designs based on the model matrix $\mathbf{X}$ and the $\mathbf{X}'\mathbf{X}$ matrix. If the design is treated as a probability measure on the design space, then design efficiencies give a measure of the information obtained per design point. If the probability measure is discrete, then $1/n$ is the weight assigned to each point, x, in the design space.

For an n-point design, the moment matrix $\mathbf{M}(\xi)$ is $\mathbf{M}(\xi) = \mathbf{X}'\mathbf{X}/n$, and the determinant of the moment matrix is $|\mathbf{M}(\xi)| = |\mathbf{X}'\mathbf{X}|/n^p$. The model: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ represents the relationship between the response Y and a set of independent factors X. $\hat{Y}$ represents the estimate of Y: $\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$. For an n-point design, a normalized measure of the variance of prediction of a point x is $d = \mathbf{x}'[\mathbf{M}(\xi)]^{-1}\mathbf{x} = n\mathbf{x}'[\mathbf{X}'\mathbf{X}]^{-1}\mathbf{x}$. If the errors are independently and identically distributed with variance σ^2, then $d\sigma^2/n$ is the variance of prediction at point $\mathbf{x}$.

The G-efficiency of a given design is defined as

$$\text{G-efficiency} = P\sigma^2/nV_{\max}(\hat{Y}) = P/d_{\max},$$

where

P is the number of parameters in the model of interest

σ^2 is the variance

n is the design size

$V_{\max}(\hat{Y}) = d_{\max}\sigma^2/n$ is the maximum prediction variance over the experimental region

$d_{\max}$ is the maximum scaled prediction variance

The G-efficiency aims at minimizing $V_{\max}(\hat{Y})$, while D-efficiency aims at maximizing $|\mathbf{M}(\xi)|$. Selection among various designs can be made by comparing their G-efficiencies. When the G-efficiency = 1.0, the design is called G-optimal. Our emphasis is on the G-efficiency criterion; additional motivation for emphasis on this criterion is given by Lucas (2007).

Kiefer and Wolfowitz (1960) proved that, when the design is a probability measure, then a G-optimal design will also be D-optimal. This equivalence does not hold for finite designs. Lucas (1976) pointed out that the best n-point design will usually have D-efficiency and G-efficiency less than 1.0. For any CRD, the value of the G-efficiency is no greater than the value of the D-efficiency. These criteria tell which is the best CRD. When each factor is independently reset on each run and the design is run in a randomized order, the maximum G-efficiency and D-efficiency that a design can achieve is 1.0. However, in split-plot experiments with proper blocking, higher efficiencies relative to a CRD can be attained.

Split-Plot Designs

Anderson and McLean (1974) indicated that, when an experiment is run as a split-plot design, restrictions on randomization occur, resulting in additional error terms.

Ju (1992) and Ju and Lucas (2002) studied split-plot designs in industrial experiments. They pointed out that running an RRO experiment with hard-to-change factors that are not reset gives the experiment a split-plot structure with the block sizes determined by the run order. Therefore, the usual assumption that the error at each combination of factors is normally distributed, with mean zero and constant variance σ^2, or $\mathbf{V} = \sigma^2[\mathbf{I}]$, will be an oversimplification. It can lead to confusing or erroneous results and can disguise the effects of the easy-to-change factors. They showed that the variance–covariance matrix $\mathbf{V}$ is

$$\mathbf{V} = \sigma_s^2[\mathbf{I}] + \sigma_w^2[\mathbf{U}\mathbf{U}'],$$

where σ_s^2 is the split-plot error, σ_w^2 is the whole-plot error, $\mathbf{U}$ is the matrix that indicates when the whole-plot error changes, and $\mathbf{U}'$ is the transpose of the $\mathbf{U}$ matrix. For a completely randomized design, $\mathbf{U} = \mathbf{I}$, so $\mathbf{V} = (\sigma_s^2 + \sigma_w^2)\mathbf{I}$. They showed that the variance of the estimates of the parameters vector $\boldsymbol{\beta}$ can be represented as

$$V(\hat{\boldsymbol{\beta}}_i) = A_i\sigma_s^2 + B_i\sigma_w^2.$$

Ju and Lucas (2002) studied the $\mathbf{V}$ matrix when the level of a single hard-to-change factor is not independently reset whenever two or more runs occur at the same level of that factor. They calculated the value of $\mathbf{V}$ over all randomizations and calculated the prediction variance of $\hat{Y}$. They analyzed several popular designs and showed that running an experiment as a split-plot design can get the proper error term, increase precision, save time, and save money. They showed that split-plotting procedures enable estimation of the whole-plot error and allow the regression coefficients of both the hard- and easy-to-change factors to be estimated more precisely. Webb et al. (2004) extended the results to L^k factorial experiments with c factors that are not reset from one run to the next, discussed a response surface example with two factors that were not reset, and presented recommendations for designing experiments containing factors that are not to be reset. In this paper, we give a catalog of optimum blocking procedures for 2^k experiments with one hard-to-change factor and give the related cost function. Bingham and Sitter (2003) noted that performing experiments in robust parameter designs as split-plot designs often provides cost savings and increased efficiency. Goos and Vandebroek (2004) pointed out the increasing popularity of split-plot designs because some of the factors under investigation in industrial experiments are often hard to change and highlighted the fact that the resulting compound symmetric error structure affects estimation and inference procedures as well as the efficiency of the experimental designs used. They computed D-optimal first and second order split-plot designs and showed that these designs, in many cases, outperform CRDs in terms of D- and G-efficiency. They suggested that split-plot designs should be considered as an alternative to CRDs even if running a CRD is affordable. Here we show our agreement with these researchers and make explicit recommendations for running 2-level factorial experiments.

Selecting Efficient Defining Relations

To obtain efficient defining relations when there is one hard-to-change factor, note that, whenever a word in the defining relation has a length that is less than or equal to the length of any term in the model

of interest, then a contribution to the whole-plot error occurs. An efficient defining relation can be selected from a list of all possible defining relations; the relation that results in the smallest prediction variance can be selected. This "brute force" approach becomes very time consuming as the number of factors increases, even for computers. Algorithms exist for the selection of defining relations for fractional factorial designs. These algorithms are discussed in the Appendix, where we modify them to allow the confounding of a hard-to-change factor with blocks.

In general, for 2^k designs in block size $2^k/2$ and 2 blocks, with one hard-to-change factor (A), the optimal block variable is A, and the optimal defining relation is $I = A$. This notation for blocking mirrors the notation used for fractional factorials. For blocking, it shows all model terms confounded with blocks; these model terms are estimated less precisely. Also in general, for 2^k designs in block size $2^k/4$ and 4 blocks, with one hard-to-change factor (A), the optimal block defining relation is I = A = interaction of all factors other than A = interaction of all factors including A.

Therefore, the word length patterns are not presented for 2^k designs with 2 or 4 blocks.

Prediction Variance and G-Efficiency

The prediction variance of $\hat{Y}$ is

$$V(\hat{Y}) = x'[X'V^{-1}X]^{-1}x.$$

For a 2^k experiment run using orthogonal blocking, the experimental region is a hypercube defined by the upper and lower levels of each factor. Because of the orthogonality and ± 1 coding used for the 2^k designs, it is easy to prove that the maximum prediction variance, $V_{\max}(\hat{Y})$, over the experimental region is

$$V_{\max}(\hat{Y}) = \frac{1}{2^k}P\sigma_s^2 + \frac{1}{2^k}P_1 b\sigma_w^2 = \frac{1}{2^k}(P\sigma_s^2 + P_1 b\sigma_w^2),$$

where:

- P is the number of parameters in the model of interest
- P_1 is the number of model terms that are in the defining relation, including I
- b is the block size
- k is the number of factors in the design
- $n = 2^k$ is the number of experimental points
- $n_b = 2^k/b$ is the number of blocks.

This formula is also the summation of the variance of each term in the model. The above equations will be used to calculate the prediction variance of $\hat{Y}$ for 2^k experiments, when there is one hard-to-change factor, for various ways of conducting the experiment. The multiplier $1/2^k$ can be omitted by considering the calculations on an information per point basis. The results for 2^3, 2^4, 2^5, 2^6, and 2^7 designs are shown in Tables 3 through 7, respectively. In these tables, the multiplier for the split-plot variance (σ_s^2) is equal to the number of parameters, P, in the model of interest. The multiplier of the whole-plot variance (σ_w^2) is $P_1 b$.

To calculate the G-efficiency of $2^k = n$ point split-plot designs on an information per point basis, recall the definition, given earlier, of the G-efficiency for

TABLE 3. Prediction Variance, G-efficiency, and Cost for 2^3 Blocked Designs

Block size b	No. of blocks n_b	Effects in the model	Variance multiplier: Split P	Variance multiplier: Whole $P_1 b$	G-efficiency relative to the CRD Based on λ value of: 0	1	10	Infinity	Cost multiplier: Hard	Easy
4	2	Main	4	8	1.00	0.67	0.52	0.50	2	8
		Main & 2 FI	7	8	1.00	0.93	0.89	0.88	2	8
		All effects	8	8	1.00	1.00	1.00	1.00	2	8
2	4	Main	4	4	1.00	1.00	1.00	1.00	4	8
		Main & 2 FI	7	6	**1.00**	**1.08**	**1.15**	**1.17**	4	8
		All effects	8	8	1.00	1.00	1.00	1.00	4	8

TABLE 4. Prediction Variance, G-efficiency, and Cost for 2^4 Blocked Designs

Block size b	No. of blocks n_b	Effects in the model	Variance multiplier: Split P	Variance multiplier: Whole $P_1 b$	G-efficiency relative to the CRD Based on λ value of: 0	1	10	Infinity	Cost multiplier: Hard	Easy
8	2	Main	5	16	1.00	0.48	0.33	0.31	2	16
		Main & 2 FI	11	16	1.00	0.81	0.71	0.69	2	16
		All effects	16	16	1.00	1.00	1.00	1.00	2	16
4	4	Main	5	8	1.00	0.77	0.65	0.63	4	16
		Main & 2 FI	11	8	**1.00**	**1.16**	**1.33**	**1.38**	4	16
		All effects	16	16	1.00	1.00	1.00	1.00	4	16
2	8	Main	5	4	**1.00**	**1.11**	**1.22**	**1.25**	8	16
		Main & 2 FI	11	10	**1.00**	**1.05**	**1.09**	**1.10**	8	16
		All effects	16	16	1.00	1.00	1.00	1.00	8	16

8 Block defining relationship: I = A = BD = CD = BC = ABD = ACD = ABC

a model with P parameters:

$$G = \frac{P}{d_{\max}} = \frac{P\sigma^2}{nV_{\max}(\hat{Y})}$$

when there are two error terms: $\sigma^2 = \sigma_s^2 + \sigma_w^2$.

Substituting the maximum prediction variance, $V_{\max}(\hat{Y})$, in the hypercube design region, gives

$$G = \frac{P(\sigma_s^2 + \sigma_w^2)}{2^k\left(\frac{1}{2^k}P\sigma_s^2 + \frac{1}{2^k}P_1 b\sigma_w^2\right)}.$$

Rearranging terms, we obtain

$$G = \frac{P\sigma_s^2 + P\sigma_w^2}{P\sigma_s^2 + P_1 b\sigma_w^2}.$$

Dividing both the numerator and denominator by $P\sigma_s^2$,

$$G = \frac{1+\lambda}{1+\frac{P_1 b}{P}\lambda},$$

where

$$\lambda = \sigma_w^2/\sigma_s^2.$$

As λ increases to infinity, the G-efficiency increases asymptotically to

$$G_{\lambda\to\infty} = P/P_1 b$$

and designs with $G_{\lambda\to\infty} > 1$, or $P > P_1 b$ are super-efficient relative to a CRD.

Using the above equations, G-efficiencies of 2^k designs can be calculated for various block sizes. In Tables 3 through 7, designs having $G > 1$, super-efficient designs, are shown with bold lettering. A super-efficient design is obtained whenever a design has a G-efficiency greater than 100% relative to a CRD. Some super-efficient designs are found for each number of factors though only a small fraction of the blocking structures shown are super efficient.

Tables 3 through 7 give the block size (b), the number of blocks (n_b), the model, the variance multipliers (P and $P_1 b$) for the split plot, and the whole-plot variance components (on a per point basis so the $2^k = n$ design size is not shown), the G-efficiency for λ values 0, 1, 10, infinity, and the cost multipliers. The cost multiplier for the hard-to-change factor is the number of blocks, and for the easy-to-change factors, it is the number of design points. The word length patterns, when there are more than 4 blocks, are shown at the bottom of the corresponding table to reflect the defining relationships.

For the main effects model, the G-efficiency always increases with number of blocks. This will hold true for all 2^k designs because of a slight extension of a result due to Fisher, who proved that a blocking relationship can always be found to estimate main effects clear of blocks with block size 2. Main effects are intentionally confounded with blocks in a split plot, so the blocking is less restrictive than the blocking considered by Fisher. Therefore, only the whole-

TABLE 5. Prediction Variance, G-efficiency, and Cost for 2^5 Blocked Designs

Block size b	No. of blocks n_b	Effects in the model	Variance multiplier: Split P	Variance multiplier: Whole $P_1 b$	G-efficiency relative to the CRD Based on λ value of: 0	1	10	Infinity	Cost multiplier: Hard	Easy
16	2	Main	6	32	1.00	0.32	0.20	0.19	2	32
		Main & 2 FI	16	32	1.00	0.67	0.52	0.50	2	32
		All effects	32	32	1.00	1.00	1.00	1.00	2	32
8	4	Main	6	16	1.00	0.55	0.40	0.38	4	32
		Main & 2 FI	16	16	1.00	1.00	1.00	1.00	4	32
		All effects	32	32	1.00	1.00	1.00	1.00	4	32
4	8	Main	6	8	1.00	0.86	0.77	0.75	8	32
		Main & 2 FI	16	12	**1.00**	**1.14**	**1.29**	**1.33**	8	32
		All effects	32	32	1.00	1.00	1.00	1.00	8	32
2	16	Main	6	4	**1.00**	**1.20**	**1.43**	**1.50**	16	32
		Main & 2 FI	16	16	1.00	1.00	1.00	1.00	16	32
		All effects	32	32	1.00	1.00	1.00	1.00	16	32

8 Block defining relationship:
I = A = BDE = CD = ABDE = ACD = BCE = ABCE
I = A = BDE = CE = ABDE = ACE = BCD = ABCD
I = A = CDE = BD = ACDE = ABD = BCE = ABCE
I = A = CDE = BE = ACDE = ABE = BCD = ABCD
Other relationships can be generated by cyclic rotation of factors.
16 Block defining relationship:
I = A = BE = CE = DE = ABE = ACE = ADE = BC = BD = CD = BCDE = ABC = ABD = ACD = ABCDE.

plot model terms will be confounded with blocks so the efficiency will increase (and $P_1 b$ will decrease) with increasing numbers of blocks. For seven and fewer factors, only the block size 2 designs are super-efficient; for eight or more factors, both block size 2 and block size 4 designs will be super-efficient. Mead (1988) pointed out a similar conclusion concerning the efficiency of incomplete block designs, although he did not address that issue for 2^k designs. All the super-efficient designs dominate CRDs because they require fewer changes of the hard-to-change factor and they have a smaller variance of prediction. Because designs with larger number of blocks require a larger number of changes for the hard-to-change factor, the experimenter will make a tradeoff between increased efficiency and increased cost.

For the main effects plus two-factor interactions model, the efficiency can reach a maximum at a certain block size and then decrease with increasing numbers of blocks. When there is an increase in the number of interactions confounded with blocks, the efficiency will decrease. Designs with large numbers of blocks are dominated by designs with fewer blocks because the fewer block designs require fewer changes of the hard-to-change factor and they are more precise. The dominated designs should seldom if ever be chosen. For the main effects plus two-factor interactions model, there is again the tradeoff between increased efficiency and increased cost up to the block size having the maximum efficiency. Again all super-efficient designs dominate CRDs, and for all numbers of factors, a super-efficient design that dominates a CRD can be found.

For the all effects model, the efficiency is 100% for all blocking structures. The completely restricted design with two blocks will be chosen if only de-

TABLE 6. Prediction Variance, G-efficiency, and Cost for 2^6 Blocked Designs

Block size b	No. of blocks n_b	Effects in the Model	Variance multiplier: Split P	Variance multiplier: Whole $P_1 b$	G-efficiency relative to the CRD Based on λ value of: 0	1	10	Infinity	Cost multiplier: Hard	Easy
32	2	Main	7	64	1.00	0.20	0.12	0.11	2	64
		Main & 2 FI	22	64	1.00	0.51	0.37	0.34	2	64
		All effects	64	64	1.00	1.00	1.00	1.00	2	64
16	4	Main	7	32	1.00	0.36	0.24	0.22	4	64
		Main & 2 FI	22	32	1.00	0.81	0.71	0.69	4	64
		All effects	64	64	1.00	1.00	1.00	1.00	4	64
8	8	Main	7	16	1.00	0.61	0.46	0.44	8	64
		Main & 2 FI	22	16	**1.00**	**1.16**	**1.33**	**1.38**	8	64
		All effects	64	64	1.00	1.00	1.00	1.00	8	64
4	16	Main	7	8	1.00	0.93	0.89	0.88	16	64
		Main & 2 FI	22	16	**1.00**	**1.16**	**1.33**	**1.38**	16	64
		All effects	64	64	1.00	1.00	1.00	1.00	16	64
2	32	Main	7	4	**1.00**	**1.27**	**1.64**	**1.75**	32	64
		Main & 2 FI	22	24	1.00	0.96	0.92	0.92	32	64
		All effects	64	64	1.00	1.00	1.00	1.00	32	64

8 Block defining relationship:
I = A = BDEF = CDE = ABDEF = ACDE = BCF = ABCF
Other relationships can be generated by making different selections and by cyclic rotation of factors.
16 Block defining relationship:
I = A = BEF = CF = DE = ABEF = ACF = ADE = BCE = BDF = CDEF = BCD = ABCE = ABDF = ACDEF = ABCD.
Other relationships can be generated by making different selections and by cyclic rotation of factors.
32 Block defining relationship:
I = A = BF = CF = DF = EF = ABF = ACF = ADF = AEF = BC = BD = BE = CD = CE = DE = CDEF = BCDE = BCEF = BCDF = BDEF = ABC = ABD = ABE = ACD = ACE = ADE = ACDEF = ABCDE = ABCEF = ABCDF = ABDEF.

sign efficiency and number of changes of the hard-to-change factor are considered. The need to estimate the whole-plot error and a desire for increased precision of the whole-plot effect could cause the designer to choose an experiment with more blocks. These issues are addressed more completely in the next sections, where we examine the cost of information.

The approach developed in this paper will be extended to fractional factorial designs elsewhere. Our approach will be compared to the literature on minimum aberration split-plot (MASP) experiments (Bingham and Sitter (1999, 2001)). Before Bingham et al. (2004), the MASP papers used a restricted definition of a split plot and only visited each level of the whole-plot factor once. This often gives designs with a higher variance of prediction than we recommend. Bingham et al. (2004) removed this restriction and tabulated many designs that are minimum variance and/or super-efficient. However, in their designs, Bingham et al. do not discuss the variance of prediction and do not always achieve the best variance of prediction. Kulahci et al. (2006) highlighted the acute need for alternatives to MASP designs.

TABLE 7. Prediction Variance, G-efficiency, and Cost for 2^7 Blocked Designs

Block size b	No. of blocks n_b	Effects in the Model	Variance multiplier: Split P	Variance multiplier: Whole $P_1 b$	G-efficiency relative to the CRD Based on λ value of: 0	1	10	Infinity	Cost multiplier: Hard	Easy
64	2	Main	8	128	1.00	0.12	0.07	0.06	2	128
		Main & 2 FI	29	128	1.00	0.37	0.24	0.23	2	128
		All effects	128	128	1.00	1.00	1.00	1.00	2	128
32	4	Main	8	64	1.00	0.22	0.14	0.13	4	128
		Main & 2 FI	29	64	1.00	0.62	0.48	0.45	4	128
		All effects	128	128	1.00	1.00	1.00	1.00	4	128
16	8	Main	8	32	1.00	0.40	0.27	0.25	8	128
		Main & 2 FI	29	32	1.00	0.95	0.91	0.91	8	128
		All effects	128	128	1.00	1.00	1.00	1.00	8	128
8	16	Main	8	16	1.00	0.67	0.52	0.50	16	128
		Main & 2 FI	29	16	**1.00**	**1.29**	**1.69**	**1.81**	16	128
		All effects	128	128	1.00	1.00	1.00	1.00	16	128
4	32	Main	8	8	1.00	1.00	1.00	1.00	32	128
		Main & 2 FI	29	20	**1.00**	**1.18**	**1.39**	**1.45**	32	128
		All effects	128	128	1.00	1.00	1.00	1.00	32	128
2	64	Main	8	4	**1.00**	**1.33**	**1.83**	**2.00**	64	128
		Main & 2 FI	29	34	1.00	0.92	0.86	0.85	64	128
		All effects	128	128	1.00	1.00	1.00	1.00	64	128

16 Block defining relationship:
I = A = BEFG = CFG = DEG = ABEFG = ACFG = ADEG
= BCE = BDF = CDEF = ABCE = ABDF = ACDEF = BCDG = ABCDG.
Other relationships can be generated by cyclic rotation of factors.
32 Block defining relationship:
I = A = BFG = CG = DF = EFG = BCF = BDG = BE = CDFG = CEF = DEG = BCD
= BCEG = CDE = BDEF = BCDEFG = ABFG = ACG = ADF = AEFG = ABCF = ABDG
= ABE = ACDFG = ACEF = ADEG = ABCD = ABCEG = ACDE = ABDEF = ABCDEFG.
Other relationships can be generated by cyclic rotation of factors.
64 Block defining relationship:
I = A = BG = CG = DG = EG = FG.
Defining relationship can be generated by multiplication.

Cost Function

The cost of running the experiment, in terms of dollars or hours, provides another criterion for selection among alternate designs. An objective of the experimental design may be to maximize the cost efficiency of the experiment. A simple cost model breaks the cost of running the experiment into two components: one for changing the level of the hard-to-change factor and the other for changing the level of the easy-to-change factors, as follows:

$$C = C_H N_H + C_E N_E = C_H n_b + C_E N_E$$

where:

C is the cost of changing the factor levels in the experiment

C_H is the cost of changing the hard-to-change factor level

$N_H = n_b$ is the number of changes of the hard-to-change factor levels including the first setting, which equals n_b, the number of blocks

C_E is the cost of changing all the easy-to-change factor levels

$N_E = n$ is the number of changes in the level of the easy-to-change factors, which equals n, the number of runs in the design.

Cost Function for Designs Run in a Random Order

The expected number of settings of the hard-to-change factor, including the initial setting, for a design run in a random order, can be obtained based on nonparametric statistics. Mood (1940) and Anbari (1993) showed that the expected number of settings, or blocks, simplifies to

$$E(\text{blocks}) = \frac{n}{2} + 1.$$

Comparison of the prediction variance, G-efficiency, and cost of designs run in a random order with designs run in blocks, reveals that, when a hard-to-change factor is present, properly blocked designs dominate designs using a randomized run order from prediction variance, G-efficiency, and cost perspectives, in each case for 2^k designs. Table 8 shows these comparisons for a 2^4 experiment.

Prediction Variance and Cost Function for Completely Randomized Designs

Using the prediction variance formula, we note that the G-efficiency for a completely randomized 2^k design is 1.0. Because the levels of the hard-to-change and easy-to-change factors are reset for each run, the number of these changes equals the number of runs, n. Therefore, the cost multiplier for all factors in these designs is equal to the design size $= n$, the highest cost alternative. For example, for a 2^4 design, the cost multiplier for all factors is 16, as shown in Table 8.

When there is a hard-to-change factor, CRDs should only be contemplated when it is particularly important to obtain precise estimates of the hard-to-change factor effect. The cost of CRDs is substantially higher than the cost of RNR designs. This explains the reason for the popularity of the latter in industry. Further discussions of this point are contained in Ju and Lucas (2002) and in Webb et al. (2004). These papers also develop the cost factor (expected number of changes) and expected variance for an RNR design. The economics of factorial and fractional factorial split-plot experiments are also discussed by Bisgaard (2000).

Table 8 compares the cost and variance multipliers for a 2^4 design using a main-effects and 2-factor interactions model. Table 8 shows that a 4-block design (with I = A = BCD = ABCD) dominates a CRD or an RNR design, and a blocked design using 8 blocks. The 4-block design's cost multiplier and the hard-to-change factor variance (σ_w^2) multiplier are both

TABLE 8. Comparison of Alternatives for the Selection of Block Size in 2^4 Design
Main-Effects and 2-Factor Interactions Model

	Variance multiplier		Cost multiplier**
	Easy (σ_s^2)	Hard (σ_w^2)	Hard
Completely randomized design***	11	11	16
Design run in a random order (RNR)	11	12*	9*
Blocked design, block size = 2	11	10	8
Blocked design, block size = 4	11	8	4
Blocked design, block size = 8	11	16	2

* Averaged over all randomizations (Anbari and Lucas (1994)).
** Expected cost = Expected number of changes.
*** In a completely randomized design (CRD), there is only one error term: $\sigma^2 = \sigma_s^2 + \sigma_w^2$.

smaller than those for the dominated design. Similar results hold for all 2^k experiments. The 2-block design is not dominated because its cost multiplier is 2. It can be worthwhile to use this design when C_H is large.

Selection of Block Size Based on Cost and Prediction Variance

A model can be constructed to minimize the cost of information obtained from the experiment as follows:

$$\text{Minimize } Z = \{\text{Cost/Information}\}$$

Information is inversely related to the error or prediction variance of $\hat{Y}$. Therefore

$$\begin{aligned}\text{Minimize } Z &= \{\text{Cost}/(1/\text{Variance})\} \\ &= \{\text{Cost} \times \text{Variance}\}.\end{aligned}$$

Therefore, minimizing of the cost of information is achieved by minimizing the product of the cost of the experiment and the prediction variance of $\hat{Y}$, as follows:

$$\text{Min } Z = C \times V(\hat{Y}),$$

where C and $V(\hat{Y})$ were given earlier. Therefore,

$$\text{Min } Z = (C_H n_b + C_E N_E) \cdot \{1/2^k \cdot (P\sigma_s^2 + P_1 b\sigma_w^2)\}$$

(remember that $n_b = N_H$ and $N_E = n$). Defining $r = C_H/C_E$ and rearranging terms, the minimization can be satisfied as

$$\text{Min } Z = (n_b r + n)(P + P_1 b\lambda).$$

The above terms can be obtained from the formulas or tables presented in this paper for various designs.

For example, for a main-effects and 2-factor interactions model for a 2^4 design, we obtain

$$Z = \begin{cases} (4r+16)(11+8\lambda) \\ \quad = 44r + 32r\lambda + 176 + 128\lambda \quad n_b = 4 \\ (2r+16)(11+16\lambda) \\ \quad = 22r + 32r\lambda + 176 + 256\lambda \quad n_b = 2. \end{cases}$$

For a main-effects plus interactions model, the 8-block design is dominated by 2- and 4-block designs, so the 8-block design need not be considered.

For given values of λ and r, the better design can be chosen. The line of indifference is found by equating both costs. Figure 2 shows the optimum block sizes and number of blocks for the main-effects plus 2-factor interactions model and Figure 3 shows the optimum block sizes and number of blocks for the main effects model. While similar figures could be drawn for other 2^k experiments, they depend on λ and r, whose values are not generally known. Conceptually, this shows that a completely restricted design with two blocks may be the best choice to minimize the cost of information when the C_H/C_E ratio (r) is large. This justifies this frequently used approach. Using n_b values above 2 but less than the smallest n_b value, having a minimum whole-plot variance multiplier ($P_1 b$) can minimize the cost of information, and it does allow the estimation of both variance components. In general, block sizes larger than 2 and up to the block size having the minimum whole-plot variance multiplier can also be considered. There is always a blocking structure that dominates a CRD and an RNR design on both a cost and variance basis, so they are not recommended. We recommend using a CRD or an RNR design much less frequently and using split-plot blocking much more frequently.

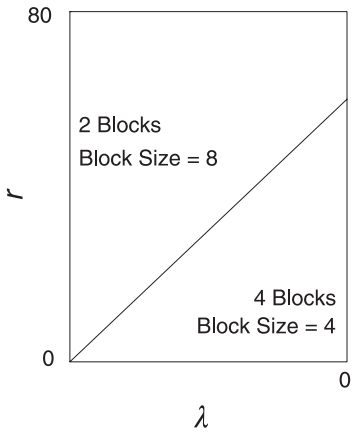


FIGURE 2. Optimum Block Size for 2^4 Experiments, Main-Effects Plus 2-Factor Interactions Model.

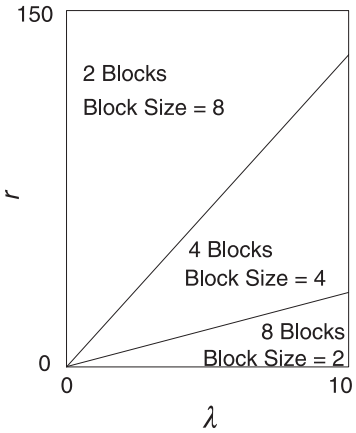


FIGURE 3. Optimum Block Size for 2^4 Experiments, Main-Effects Model.

Conclusions

In this paper, we showed how to design and run 2^k experiments to get the smallest prediction variance and lowest cost. We presented and demonstrated an algorithm to allow the selection of efficient defining relations in 2^k designs where there is one hard-to-change factor. We presented methods for calculating the prediction variance and G-efficiency when there are hard-to-change and easy-to-change factors and demonstrated the calculations by applying them to 2^3, 2^4, 2^5, 2^6, and 2^7 designs and by tabulating the results for various block sizes.

We developed a cost function for the experimental design, constructed a cost model, and found the block size that minimizes the cost of information for the model of interest. We compared variance and cost results to those of CRDs and RNR designs. We showed that properly blocked designs dominate the current practice of running experiments in a random order from prediction variance (G-efficiency) and cost perspectives. It may be appropriate to include these results in books for design of experiments, as they demonstrate the superiority of blocked split-plot experiments over randomized designs. Future work will extend the results presented here to designs that have two or more hard-to-change factors.

Appendix
Algorithm for Selecting Efficient Defining Relationships

Development of the Algorithm

Greenfield (1976, 1978), Franklin and Bailey (1977), and Franklin (1985) proposed algorithms for the selection of defining relations in various experimental designs. Anbari (1993, 2004) discussed the existing algorithms and suggested modifications to support the selection of design generators that allow the confounding of hard-to-change factor(s) with blocks and that result in designs having the minimum prediction variance of $\hat{Y}$ for the models of interest: main effects, main effects plus 2-factor interactions, and all effects. Loeppky et al. (2006) discussed the construction of optimal nonregular fractional factorial designs with two different types of factors based on the word-length pattern to emphasize the estimation of the effects of interest.

The algorithm presented in this paper requires the confounding of the hard-to-change factor and seeks the defining relation(s) that have the smallest number of words that affect the prediction variance of the main effects model and the main-effects and 2-factor interactions model. The algorithm is based on seeking words of length 3 or more because these words do not contribute to the prediction variance of the main effects model or the main-effects and 2-factor interactions model. Such words have an impact only on models having 3 or more factor interactions. Because all words in the defining relation have a potential impact on the prediction variance, it is essential that as many words as possible in the defining relation and its generalized interactions be of length 3 or more.

The proposed algorithm accomplishes this objective by requiring 3 independent sources for letters to be used in the defining relation generator. By placing different letters at the top of the table used in the algorithm and selecting words for the defining relation generator from 2 different columns, the generalized interaction resulting from the multiplication of any 2 words selected from any 2 different columns has at least 2 letters. By further placing other letters and their generalized interactions at the beginning of each row in the table used in the algorithm and selecting words for the defining relation generator from 2 different rows, the generalized interaction resulting from the multiplication of any 2 words selected from any 2 different rows has at least 1 additional letter. Thus, the resulting generalized interactions selected in this fashion will have at least 3 letters and will not impact the prediction variance of the main effects model or the main-effects and 2-factor interactions model.

The words used for the defining relation generator are selected to be the longest available words from each column in available rows. Therefore, these words, and the entire defining relation, will have the least impact on the prediction variance of the models under consideration.

Steps of the Algorithm

The algorithm is carried out in the following steps:

1. For a 2^k design with block size of 2^p (and 2^{k-p} blocks), the number of independent generators required is $k - p$. Therefore, construct a table having the number of columns equal to $k - p$.
2. List the hard-to-change factor(s) at the top of the table, followed by as many factor(s) as needed to make up the $k - p$ generators.
3. List the remaining factors at the left side of the table, starting with I. Complete the column by showing all generalized interactions of these factors.

4. Multiply the column created in step 3 by the heading of each column and list the resulting words in the appropriate column. For I, the only eligible results are those resulting from multiplying I by the hard-to-change factor(s). For the hard-to-change factor(s), the only eligible results are those resulting from multiplying the hard-to-change factor(s) by I.
5. Start to construct the defining relation by selecting I and the hard-to-change factor(s).
6. Select the longest word from the first column and include it in the defining relation. Eliminate that column and corresponding row from further consideration.
7. Select the longest word from the next column and include it in the defining relation. Do not use a word from a previously used row. If all rows have been used, select the longest word(s) in the next column(s) and include it(them) in the defining relation.
8. Stop when the defining relation has in it the required number of independent generators. This should coincide with reaching the last column in the table. Generate the balance of the defining relation by multiplication.
9. If the rows were exhausted at the same time as the columns (Example: 2^6 design in block size = 4) or prior to the columns (example: 2^7 design in block size = 4), then an efficient form of the defining relation has been obtained. "Efficient" is used here to refer to a defining relation that has a small number of words that affect the prediction variance of the model of interest. If the defining relation has the smallest possible number of words that affect the prediction variance of the model(s) of interest, then this defining relation is the most efficient and can be referred to as the optimal defining relation.
10. If the columns were exhausted before the rows (example: 2^5 design in block size = 4), then other form(s) of an efficient defining relation exists, which may be equivalent to the form already obtained. The other form(s) can be generated by selecting the longest word in a different column and proceeding as described above to obtain another efficient form of the defining relation.

The following steps allow finding other defining relations by enumeration, but do not generate new forms of the defining relation:

11. Rotate the factors cyclically by moving the last factor at the top of the table to the bottom of the factors in the column at the left of the table. Move the first factor after I from that column and place it at the top of the table, after the hard-to-change factor(s). Do not rotate factors if the table contains only one row other than I or one column other than the hard-to-change factor(s). The defining relation form would have already been found in these cases.
12. Repeat steps 3 through 10 for the new table.
13. Repeat steps 11 and 12 until the factor that originally appeared at the top of the table in the first column after the hard-to-change factor(s) reaches the row immediately below I in the column at the left of the table.
14. From the generated list of defining relations, select the one that has the smallest number of words that affect the prediction variance of the model(s) of interest. This can be referred to as the optimal defining relation.

Example of the Algorithm

An example of applying the algorithm to a 2^5 design in block size = 4 is shown in Table A.

TABLE A: 2^5 Design, Block Size = 4

	A	B	C
I	A	—	—
D	—	BD	CD
E	—	BE	CE
DE	—	BDE	CDE

Block size = $4 = 2^2$, number of blocks = $2^{5-2} = 2^3 = 8$, number of columns at top of table = 3.
Optimal defining relations:
(1) I = A = BDE = CD
= ABDE = ACD = BCE = ABCE
(2) I = A = BDE = CE
= ABDE = ACE = BCD = ABCD
(3) I = A = CDE = BD
= ACDE = ABD = BCE = ABCE
(4) I = A = CDE = BE
= ACDE = ABE = BCD = ABCD
Generate other relations by cyclic rotation of factors.

References

Anbari, F. T. (1993). "Experimental Designs for Quality Improvement When There Are Hard-to-Change Factors and Easy-to-Change Factors". Ph.D. Thesis, Drexel University, Philadelphia, PA.

Anbari, F. T. (2004). "Optimal Blocking with One Hard-to-Change Factor". *Proceedings of the American Statistical Association Joint Statistical Meetings*, Toronto, Ontario, Canada.

Anbari, F. T. and Lucas, J. M. (1994). "Super-Efficient Designs: How to Run Your Experiment for Higher Efficiency and Lower Cost". *Proceedings, ASQ 48th Annual Quality Congress*.

Anderson, V. L. and McLean, R. A. (1974). "Restriction Errors: Another Dimension in Teaching Experimental Statistics". *The American Statistician* 28, pp. 145–152.

Bingham, D. and Sitter, R. R. (1999). "Minimum-Aberration Two-Level Fractional Factorial Split-Plot Designs". *Technometrics* 41, 62–70.

Bingham, D. R. and Sitter, R. R. (2001). "Design Issues in Fractional Factorial Split-Plot Experiments". *Journal of Quality Technology* 33, pp. 2–15.

Bingham, D. R. and Sitter, R. R. (2003). "Fractional Factorial Split-Plot Designs for Robust Parameter Experiments". *Technometrics* 45, pp. 80–89.

Bingham, D. R.; Schoen, E. D.; and Sitter, R. R. (2004). "Designing Fractional Factorial Split-Plot Experiments with Few Whole-Plot Factors". *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 53, pp. 325–339.

Bisgaard, S. (2000). "The Design and Analysis of $2^{k-p} \times 2^{q-r}$ Split Plot Experiments". *Journal of Quality Technology* 32, pp. 39–56.

Franklin, M. F. (1985). "Selecting Defining Contrasts and Confounded Effects in p^{n-m} Factorial Experiments". *Technometrics* 27, pp. 165–172.

Franklin, M. F. and Bailey, R. A. (1977). "Selection of Defining Contrasts and Confounded Effects in Two-level Experiments". *Applied Statistics* 26, pp. 321–326.

Ganju, J. and Lucas, J. M. (1997). "Bias in Test Statistics when Restrictions in Randomization Are Caused by Factors". *Communications in Statistics: Theory and Methods* 26, pp. 47–63.

Ganju, J. and Lucas, J. M. (1999). "Detecting Randomization Restrictions Caused by Factors". *Journal of Statistical Planning and Inference* 81, pp. 129–140.

Ganju, J. and Lucas, J. M. (2005). "Randomized and Random Run Order Experiments". *Journal of Statistical Planning and Inference* 133, pp. 199–210.

Goos, P. and Vandebroek, M. (2004). "Outperforming Completely Randomized Designs". *Journal of Quality Technology* 36, pp. 12–26.

Goos, P.; Langhans, I.; and Vandebroek, M. (2006). "Practical Inference from Industrial Split-Plot Designs". *Journal of Quality Technology* 38, pp. 163–179.

Greenfield, A. A. (1976). "Selection of Defining Contrasts in Two-Level Experiments". *Applied Statistics* 25, pp. 64–67.

Greenfield, A. A. (1978). "Selection of Defining Contrasts in Two-level Experiments—A Modification". *Applied Statistics* 27, p. 78.

Ju, H. L. (1992). "Split Plotting and Randomization in Industrial Experiments". Ph.D. Dissertation, University of Delaware, Newark, DE.

Ju, H. L. and Lucas, J. M. (2002). "L^k Factorial Experiments with Hard-to-Change Factors and Easy-to-Change Factors". *Journal of Quality Technology* 34, pp. 411–421.

Kiefer, J. (1959). "Optimum Experimental Designs". *Journal of the Royal Statistical Society, Series B* 21, pp. 272–319.

Kiefer, J. (1961). "Optimum Designs in Regression Problems II". *The Annals of Mathematical Statistics* 32, pp. 298–325.

Kiefer, J. and Wolfowitz, J. (1960). "The Equivalence of Two Extremum Problems". *Canadian Journal of Mathematics* 12, pp. 363–366.

Kulahci, M.; Ramírez, J. G.; and Tobias, R. (2006). "Split-Plot Fractional Designs: Is Minimum Aberration Enough?" *Journal of Quality Technology* 38, pp. 56–64.

Loeppky, J. L.; Bingham, D.; and Sitter, R. R. (2006). "Constructing Non-Regular Robust Parameter Designs". *Journal of Statistical Planning and Inference* 136, pp. 3710–3729.

Lucas, J. M. (1976). "Which Response Surface Design is Best". *Technometrics* 18, pp. 411–417.

Lucas, J. M. (1978). "Discussion of D-Optimal Fractions of Three-Level Factorial Designs". *Technometrics* 20, pp. 381–382.

Lucas, J. M. (2007). "Letters to the Editor: Comments on Optimal Designs for Second Order Polynomial Models (*JQT*, 2005, 37, pp. 253–266)". *Journal of Quality Technology* 39, pp. 90–91.

Lucas, J. M. and Ju, H. L. (1992). "Split Plotting and Randomization in Industrial Experiments". *Transactions, ASQ Annual Quality Congress*, pp. 372–384.

Meade, R. (1988). *The Design of Experiments: Statistical Principles for Practical Applications*. Cambridge University Press, Cambridge, UK.

Mood, A. M. (1940). "The Distribution Theory of Runs". *Annals of Mathematical Statistics* 11, pp. 367–392.

O'Neill, J. and DeBrunce, V. (2001). "Robust Product Design to Minimize Counted Defects". *Statistical Methods in Quality: In Celebration of National Quality Month*, Philadelphia Section of ASQ and Penn State Great Valley.

Taguchi, G. (1984). *Quality Engineering: Methodology & Application*. American Supplier Institute, Romulus, MI.

Taguchi, G. (1987). *System of Experimental Design*, Vol. 1 and 2. UNIPUB/Kraus International Publications, White Plains, NY, and American Supplier Institute, Dearborn, MI.

Webb, D. F.; Lucas, J. M.; and Borkowski, J. J. (2004). "Factorial Experiments when Factor Levels Are Not Necessarily Reset". *Journal of Quality Technology* 36, pp. 1–11.

Wortham, A. W. and Smith, T. E. (1959). *Practical Statistics in Experimental Design: Methods/Applications*. Dallas Publishing House, Dallas, TX.

Youden, W. J. (1972). "Enduring Values". *Technometrics* 14, pp. 1–11.