The 2^{k-p} Fractional Factorial Designs* Part I.

G. E. P. BOX AND J. S. HUNTER

Statistics Department, University of Wisconsin and Mathematics Research Center, University of Wisconsin

"Cats is dogs and dogs is dogs and rabbits is dogs, and squirrels in cages is parrots . . . ."

1: THE TWO-VERSION FACTORIALS AND FRACTIONALS

A full 2^k factorial design requires all combinations of two versions of each of k variables. If a variable is continuous, the two versions become the high and low level of that variable. If a variable is qualitative the two versions correspond to two types, sometimes the presence and absence of the variable.

The runs comprising the experimental design are conveniently set out in either of two notations as illustrated for the eight runs comprising a 2^3 factorial in Table 1.

TABLE 1
Alternative Notations for the 2^3 Factorial Design

Run Number	*Notation* 1 Variables $A\ B\ C$	*Notation* 2 Variables **1 2 3**
1	1	− − −
2	a	+ − −
3	b	− + −
4	ab	+ + −
5	c	− − +
6	ac	+ − +
7	bc	− + +
8	abc	+ + +

In the first notation the variables are identified by capital letters, and their two versions by the presence or absence of the corresponding lower case letter. When all the variables are at their "low" level or version a "1" is used. In the second notation the variables are identified by numbers and the two versions of each variable by either a minus and plus sign, or by minus and plus one. The experimental design can then be viewed geometrically. A run is represented by a point whose coordinates are the ± 1 versions for that run. For example,

* Sponsored by the United States Army under Contract No. DA-11-022-ORD-2059.

the 2^3 factorial will provide the eight vertices of a cube in a three-dimensional coordinate system. The notation using minus and plus signs is used in this paper. The list of experimental runs is called the *design matrix* and is denoted by **D**. For a 2^k factorial, the design matrix contains k columns and $N = 2^k$ rows. There is a column for each of the k variables, and each row gives the combination of versions for each run.

In Table 1 the runs are listed in *standard order*. The elements of the first column are alternate minus and plus signs. The elements of the second column are alternate pairs of minus and plus signs, the elements of the third column alternate groups of four minus and plus signs and so on. The last column consists of 2^{k-1} minus signs followed by 2^{k-1} plus signs.

The Estimates

On the assumptions that the observations are uncorrelated and have equal variance, then the 2^k factorial designs provide independent minimum variance estimates of the grand average and of the 2^{k-1} effects:

$$\begin{array}{ll} k & \text{main effects,} \\ \dfrac{k(k-1)}{2} & \text{two-factor interaction effects,} \\ \dfrac{k(k-1)(k-2)}{2\cdot 3} & \text{three-factor interaction effects,} \\ \vdots & \\ \dfrac{k(k-1)(k-2)\cdots(k-h-1)}{h!} & h\text{-factor interaction effects,} \end{array} \tag{1}$$

and finally a single k-factor interaction effect.

Although the Yates' Algorithm (1) provides a quick method for calculating these estimates, a longer, but more basic calculation technique will now be described. In Table 2, where for convenience a 2^3 design is used, a *matrix of independent variables* **X** is generated from the design matrix **D**. In general the individual elements for an ij interaction column in **X** are obtained by multi-

TABLE 2

Design Matrix D			Matrix of Independent Variables X								Observations Y
1	2	3	I	1	2	3	12	13	23	123	
−	−	−	+	−	−	−	+	+	+	−	2
+	−	−	+	+	−	−	−	−	+	+	10
−	+	−	+	−	+	−	−	+	−	+	8
+	+	−	+	+	+	−	+	−	−	−	12
−	−	+	+	−	−	+	+	−	−	+	6
+	−	+	+	+	−	+	−	+	−	−	8
−	+	+	+	−	+	+	−	−	+	−	6
+	+	+	+	+	+	+	+	+	+	+	4

plying the corresponding elements of the separate i and j columns. Similarly, the elements of the ijk interaction column are given by the product of the elements of the columns labelled ij and k and so on. The first column of **X** consists entirely of plus signs and is used to provide an estimate of the mean. For a 2^k design the full matrix of independent variables **X** contains 2^k columns as well as 2^k rows. The elements of the column **Y** in Table 2 are the observations recorded at each of the 2^k experiments. The estimate of the effect $\mathbf{ij} \cdots \mathbf{k}$ is obtained by taking the sum of products between the elements of **Y** and the corresponding elements of the column $ij \cdots k$ and dividing this product by $N/2$ where $N = 2^k$, e.g.,

$$\mathbf{ij} \cdots \mathbf{k} \text{ effect} = \frac{2}{\mathrm{N}} \sum y\{i.j. \cdots. k\} \tag{2}$$

where $\{i.j. \cdots. k\}$ stands for the elements of the $ij \cdots k$ column and the summation is taken over all N products. Thus, using the data from Table 2 the **1 3** interaction effect is

$$\mathbf{1\ 3} = \tfrac{1}{4}(2 - 10 + 8 - 12 - 6 + 8 - 6 + 4) = -3.0.$$

where here and henceforth numerals appearing in bold type are used to identify the main effects and interactions. Solving for all the effects gives:

Main Effects	Two-factor Interactions	Three-factor Interaction
1 = 3.0	**1 2** = −2.0	**1 2 3** = zero
2 = 1.0	**1 3** = −3.0	
3 = −2.0	**2 3** = −3.0	

Each estimated effect has variance

$$\text{Variance (effect)} = 4\sigma^2/N \tag{3}$$

where σ^2 is the variance of the individual observations.

The average is obtained by taking the sum of products of the column **I** with the observation column **Y** and dividing the result by N, thus

$$\text{average} = \bar{y} = \sum y\{\mathbf{I}\}/N. \tag{4}$$

Thus $\bar{y} = 56/8 = 7.0$ with variance σ^2/N. By this process 2^k estimates can be obtained from 2^k runs and when k is large the wealth of such estimates becomes almost an embarrassment. However, in many practical situations, the higher order interaction effects can often be hopefully supposed to be negligible in size. For example, with continuous variables it is reasonable to expect the response to vary *smoothly*. When factorial designs are correctly used to study qualitative variables it is because certain aspects of *similarity* are expected in the responses at the different versions. Thus, two solvents and two differently shaped particles may with profit be studied in a factorial design when at least some aspect of similarity in behavior of these variables might be expected.

In the conditions of smoothness and similarity commonly encountered, the three-factor and multi-factor interaction effects are often negligible. When this is the case, fractional designs using a smaller number of runs may be employed for although in these fractional designs the effects of the major interest are confused with higher order effects, nevertheless, the latter are small enough to be ignored. In some situations the total number of variables k is large, but only a few (say $p = 2$ or 3) are expected to have any effect. In this situation designs which are fractional in the k variables may be chosen which have the property that they are complete factorials in any sub-group of p variables.

For illustration, we first discuss the one half fraction of the 2^4 design.

One-half Fraction of the 2^4 Factorial

Since the design is to contain $2^{4-1} = 8$ runs a 2^3 factorial design is first written down. The $-$ and $+$ elements associated with the **1 2 3** interaction column then are used to identify the $-$ and $+$ versions of variable **4**. The resulting eight combinations shown in Table 3 give a particular half replicate or "fractional" of the complete 2^4 design. A $(\frac{1}{2})^p$ fraction of a 2^k factorial design is called a 2^{k-p} *fractional*, or more exactly, a 2^{k-p} fractional factorial. The present design is therefore a 2^{4-1} fractional.

TABLE 3
Constructing the 2^{4-1} Fractional Factorial Design

Design Matrix				Observations
1	**2**	**3**	**1 2 3 = 4**	Y
$-$	$-$	$-$	$-$	8.7
$+$	$-$	$-$	$+$	15.1
$-$	$+$	$-$	$+$	9.7
$+$	$+$	$-$	$-$	11.3
$-$	$-$	$+$	$+$	14.7
$+$	$-$	$+$	$-$	22.3
$-$	$+$	$+$	$-$	16.1
$+$	$+$	$+$	$+$	22.1

With a full 2^4 design, sixteen effects can be estimated: the grand average, four main effects, six two-factor interactions, four three-factor interactions and a single four factor interaction. With only eight observations it is clearly impossible to obtain sixteen independent estimates. We note that the combination of observations used to estimate the main effect **4** is identical to that used to estimate the three-factor interaction effect **1 2 3**. The estimates of **4** and **1 2 3** are said to be *confounded*. The "**4**" effect really estimates the *sum* of the effects of **4** and **1 2 3**.

Study of Table 3 will show that other estimates such as **1 2** and **3 4** are also confounded. It is desirable to have a general method which enables one to determine which effects are confounded. This is accomplished for this design by inducing the equality **4** = **1 2 3** where the multiplication product **1 2 3** refers to the multiplication of the individual elements in the corresponding columns **1**, **2** and **3**. Now it is obvious that by multiplying the elements in any column by a column of identical elements, a column of plus signs will result. Since a column of plus signs corresponds to **I** we have $\mathbf{1} \times \mathbf{1} = \mathbf{1}^2 = \mathbf{I}$ and similarly that $\mathbf{2}^2 = \mathbf{I}$, $\mathbf{3}^2 = \mathbf{I}$

and $\mathbf{4}^2 = \mathbf{I}$. This identity supplies the key to the remaining relationships. On multiplying both sides of the equation $\mathbf{4} = \mathbf{1\ 2\ 3}$ by $\mathbf{4}$ we get

$$\mathbf{4}^2 = \mathbf{1\ 2\ 3\ 4} \qquad \text{that is} \qquad \mathbf{I} = \mathbf{1\ 2\ 3\ 4}. \tag{5}$$

This identity is readily confirmed for if the elements in columns **1**, **2**, **3** and **4** are multiplied together we obtain a column of plus signs, that is **I**.

The interaction **1 2 3 4** associated with **I** is said to be a *generator* of the design. In this particular instance there is only one generator so this provides the *defining relation* $\mathbf{I} = \mathbf{1\ 2\ 3\ 4}$ which is the key to all the relationships which exist between the effects.

Aliases and Linear Combinations of Effects

Suppose we wish to know which effect is confounded with the main effect **3**. Multiplying both sides of the defining relation by **3** gives $\mathbf{3} = \mathbf{1\ 2\ 3}^2\ \mathbf{4} = \mathbf{1\ 2\ I\ 4} = \mathbf{1\ 2\ 4}$ since multiplication by **I** (a column of plus signs) leaves the elements in any column unchanged. Thus, the main effect **3** is confounded with the three-factor interaction **1 2 4**. Similarly, we find that the two-factor interaction **3 4** is confounded with **1 2** and so on. The quantities so associated are called *aliases*. If we now proceed to estimate the main effect **3** we will in fact obtain the *sum* of the estimates of the main effect **3** and the three-factor interaction **1 2 4**. The estimate of **3** is really an estimate of the combination of the effects $\mathbf{3} + \mathbf{1\ 2\ 4}$. Eight linear combinations of effects ℓ_I, ℓ_1, ℓ_2, $\cdots$ are available. Thus ℓ_1, $= \frac{1}{4} \sum y\{\mathbf{1}\}$ or equally $\ell_1 = \frac{1}{4} \sum y\{\mathbf{2\ 3\ 4}\}$. Similarly $\ell_{12} = \frac{1}{4} \sum y\{\mathbf{1\ 2}\}$ or equally $\ell_{12} = \frac{1}{4} \sum y\{\mathbf{3\ 4}\}$. Using the defining relation we find that these linear combinations estimate the quantities given in Table 4, the subscript on the ℓ's identifying the first effect in the linear combination.

TABLE 4

$\ell_I = \text{average} + \mathbf{1\ 2\ 3\ 4}$	$\ell_4 = \mathbf{4} + \mathbf{1\ 2\ 3}$
$\ell_1 = \mathbf{1} + \mathbf{2\ 3\ 4}$	$\ell_{12} = \mathbf{1\ 2} + \mathbf{3\ 4}$
$\ell_2 = \mathbf{2} + \mathbf{1\ 3\ 4}$	$\ell_{13} = \mathbf{1\ 3} + \mathbf{2\ 4}$
$\ell_3 = \mathbf{3} + \mathbf{1\ 2\ 4}$	$\ell_{14} = \mathbf{1\ 4} + \mathbf{2\ 3}$

The variance of these estimates is $\sigma^2/2$. The average $\bar{y}$ has variance $\sigma^2/8$.

On studying Table 4 we see that the two-factor interactions are mutually

TABLE 5

The eight linear combinations of effects from a 2^{4-1} design with defining relation $\mathbf{I} = \mathbf{1\ 2\ 3\ 4}$

$\ell_I = \text{average} + \mathbf{1\ 2\ 3\ 4} =$	15.0	$\ell_4 = \mathbf{4} + \mathbf{1\ 2\ 3} =$	0.8
$\ell_1 = \mathbf{1} + \mathbf{2\ 3\ 4} =$	5.4	$\ell_{12} = \mathbf{1\ 2} + \mathbf{3\ 4} =$	−1.6
$\ell_2 = \mathbf{2} + \mathbf{1\ 3\ 4} =$	−0.4	$\ell_{13} = \mathbf{1\ 3} + \mathbf{2\ 4} =$	1.4
$\ell_3 = \mathbf{3} + \mathbf{1\ 2\ 4} =$	7.6	$\ell_{14} = \mathbf{1\ 4} + \mathbf{2\ 3} =$	1.0

confounded in pairs, but assuming that the three and four factor interactions are either non-existent or negligible the estimates ℓ_I , ℓ_1 , ℓ_2 , ℓ_3 and ℓ_4 can be taken to be estimates of the average and the main effects **1, 2, 3** and **4.** If, furthermore, prior knowledge is available that, for example, the **3 4** interaction effect was negligible, then the estimate ℓ_{12} could be taken to estimate the **1 2** interaction effect alone.

The Alternative Fraction

In the above example, in forming the 2^{4-1} design, the factor **4** was associated with the three-factor interaction **1 2 3.** In standard ordering, the elements of the three-factor interaction column, and hence of factor **4**, are

$$- \quad + \quad + \quad - \quad + \quad - \quad - \quad +.$$

The factor **4** can either use these elements as they stand, or it can be associated with the negative of the **1 2 3** effect, that is, with the elements

$$+ \quad - \quad - \quad + \quad - \quad + \quad + \quad -.$$

In the first case $\mathbf{4} = \mathbf{1\ 2\ 3}$ that is $\mathbf{I} = \mathbf{1\ 2\ 3\ 4}$, and in the second case $-\mathbf{4} = \mathbf{1\ 2\ 3}$ that is $\mathbf{I} = -\mathbf{1\ 2\ 3\ 4}$. The designs for these two 2^{4-1} fractional factorials are given in Table 6. The two parts together constitute a complete 2^4 factorial design.

TABLE 6

The Design Matrices for the two 2^{4-1} Fractional Factorials with Defining Relations $\mathbf{I} = \mathbf{1\ 2\ 3\ 4}$ *and* $\mathbf{I} = -\mathbf{1\ 2\ 3\ 4}$.

Defining Relation I = 1 2 3 4				Observations	Defining Relation I = −1 2 3 4				Observations
1	**2**	**3**	**4**		**1**	**2**	**3**	**4**	
−1	−1	−1	−1	8.7	−1	−1	−1	1	11.8
1	−1	−1	1	15.1	1	−1	−1	−1	13.6
−1	1	−1	1	9.7	−1	1	−1	−1	9.2
1	1	−1	−1	11.3	1	1	−1	1	14.6
−1	−1	1	1	14.7	−1	−1	1	−1	15.8
1	−1	1	−1	22.3	1	−1	1	1	24.0
−1	1	1	−1	16.1	−1	1	1	1	16.4
1	1	1	1	22.1	1	1	1	−1	24.2

Table 6 shows a further set of observations associated with the second fraction. In Table 7 eight linear combinations of effects ℓ'_I , ℓ'_1 , ℓ'_2 , $\cdots$ associated with the fraction having defining relation $\mathbf{I} = -\mathbf{1\ 2\ 3\ 4}$ are given. If both fractions are present, then simple addition and subtraction of the ℓ and ℓ' linear combinations will provide unconfounded estimates of all the effects. For example, the main effect **1**, unconfounded with the **2 3 4** interaction is given by $\frac{1}{2}(\ell_1 + \ell'_1) = 5.6$. Similarly, the **2 3 4** interaction unconfounded by the main effect **1** is obtained from $\frac{1}{2}(\ell_1 - \ell'_1) = -0.20$. The average response, when both fractions are present, is given by $\frac{1}{2}(\ell'_I + \ell_I) = 15.6$.

The estimates obtained by taking the sums and differences of the linear

TABLE 7

The eight linear combinations of effects from a 2^{4-1} design with defining relation I = −1 2 3 4

ℓ'_I = average − **1 2 3 4** = 16.2	ℓ'_4 = **4** − **1 2 3** = −1.0
ℓ'_1 = **1** − **2 3 4** = 5.8	ℓ'_{12} = **1 2** − **3 4** = 0.8
ℓ'_2 = **2** − **1 3 4** = −0.2	ℓ'_{13} = **1 3** − **2 4** = 2.2
ℓ'_3 = **3** − **1 2 4** = 7.8	ℓ'_{14} = **1 4** − **2 3** = 0.6

combinations computed from the individual fractional factorials are the same as would be obtained from an analysis of a full 2^4 design.

The $\frac{1}{2}$ Fractions of the 2^k Designs

Any interaction or main effect can be used to split a full 2^k factorial into two half fractions. However, given the assumptions that the higher the order of the interaction the less likely the effect is to occur, there is clearly an advantage in using the interaction of highest order to make the split. The generator is then **1 2 3** ⋯ **k** and the defining relation **I** = **1 2 3** ⋯ **k**.

The $\frac{1}{2}$ fractions of all the 2^k factorial designs are best obtained by first writing down the design matrix for a full 2^{k-1} factorial and then adding the kth variable by identifying its + and − versions with the + and − signs of the highest order interaction **1 2 3** ⋯ (**k** − 1). Thus the 2^{3-1} factorial is constructed by writing down the design matrix for the 2^2 factorial and then equating variable **3** with the **1 2** interaction. Similarly, the 2^{5-1} factorial is given by writing down the sixteen runs of the 2^4 and then equating the signs of variable **5** with the signs of the **1 2 3 4** interaction. The defining relations for these $\frac{1}{2}$ replicate designs are thus

Design	Defining Relations	
2^{3-1}	**I** = **1 2 3**	
2^{4-1}	**I** = **1 2 3 4**	(6)
2^{5-1}	**I** = **1 2 3 4 5**	

The extension to the half-replicate designs for $k > 5$ is obvious. However, for $k > 5$ these half-replicate designs permit the estimation of a plethora of linear combinations of effects, many of which are combinations of higher order interactions solely. We are therefore interested in still smaller fractions of the 2^k designs, that is, in the 2^{k-p} fractional factorials for $p > 1$. For such designs there is not one, but p generators which combine to provide the defining relation. Before discussing these designs, it is profitable first to discuss their areas of application.

2: AREAS OF APPLICATION

Fractional designs are of value in a number of different circumstances:

1) where certain interactions can be assumed non-existent from prior knowledge,

2) in "screening" situations where it is expected that the effects of all but a few of the variables studied will be negligible,
3) where groups of experiments are run in sequence and ambiguities remaining at a given stage of experimentation can be resolved by later groups of experiments,
4) where certain variables, which may interact, are to be studied simultaneously with other variables whose influence, if any, can be described by main effects only.

Some Interactions Non-Existent, A Priori

As already noted, when properties of smoothness and similarity exist, interactions between three or more variables are often negligible. In addition, the physical nature of a problem is sometimes such that certain interactions must be small or non-existent. In these circumstances we can then use arrangements in which the effects expected to be real are confounded only with interactions expected to be negligible. For example, in Table 3 the estimate of the **1 2 3** interaction effect is perfectly confounded with the main effect **4**. Under the assumption that the three-factor interaction is small, the estimate can be taken as the main effect of **4** alone.

In most practical situations, to say that we *assume*, a priori, that certain effects are negligible would be too strong. Frequently, limitations of time and money do not allow the luxury of the certainty obtainable from exploring an entirely comprehensive model which allows for every contingency. We *tentatively entertain* the possibility of negligible interactions and try to check assumptions as the evidence unfolds.

Screening Situations

Situations often occur where not very much is known about the variables that influence some response. Any subset of a large number of variables might be important, but *which* variables form this subset is unknown. Although usually the number of variables under study will be greater than four, the application of fractionals to this situation can be illustrated with the 2^{4-1} design given in Table 3. It will be seen that if any one variable out of the four produces a large effect, then no matter which variable it is, the design may be regarded as a 2^1 factorial replicated four times in the important variable. If any two variables are producing large effects, the design becomes a full 2^2 factorial replicated twice in these variables. If any three variables are producing large effects, again the design becomes a full 2^3 factorial in these variables. Fractionals for use in screening situations which are replicated factorials for any number up to three variables out of sixteen can be obtained using only thirty-two runs. For picking out the two or three important variables from among a large group of variables, these designs are very useful.

Sequential Groups of Experiments

Fractional factorials are of considerable value in the common situation where experiments are performed in sequence. Having performed one fraction, the results can be reviewed and where there is ambiguity due to the confounding of particular estimates, or experimental error, a further group of experiments can be selected to resolve the uncertainty.

Simultaneous Study of "Major" and "Minor" Variables

It sometimes happens that there exists a group of "major" variables whose study is the chief objective of the investigation. In addition there may be a number of "minor" variables which are expected to have negligible effects. Fractional designs are available in which both kinds of variables are included simultaneously, the main effects and interactions of the major variables estimated without bias, and the main effects of the minor variables checked. The assumption made is that interactions between the minor variables will be negligible.

3: Special Types of 2^k Factorials

Fractional factorial designs can, for convenience, be divided into types. In general the higher the degree of fractionation the more comprehensive the assumptions needed to make unequivocal interpretation possible. The following three types of designs are discussed:

(i) Designs of Resolution III in which no main effect is confounded with any other main effect, but main effects are confounded with two-factor interactions and two-factor interactions with one another. The 2^{3-1} design is of Resolution III.
(ii) Designs of Resolution IV in which no main effect is confounded with any other main effect *or* two-factor interaction, but where two-factor interactions are confounded with one another. The 2^{4-1} design is of Resolution IV.
(iii) Designs of Resolution V in which no main effect or two-factor interaction is confounded with any other main effect or two-factor interaction but two factor interactions are confounded with three factor interactions. The 2^{5-1} design is of Resolution V.

In general, a design of resolution R is one in which no p factor effect is confounded with any other effect containing less than $R - p$ factors.

To identify the resolution of a fractional factorial design, the appropriate Roman numeral subscript is used. Thus, rewriting Equation (6) along with the defining relations for both one-half functions we have

$$\begin{array}{ll} \text{Design} & \text{Defining Relations} \\ 2^{3-1}_{\text{III}} & \mathbf{I} = \pm\mathbf{1\ 2\ 3} \\ 2^{4-1}_{\text{IV}} & \mathbf{I} = \pm\mathbf{1\ 2\ 3\ 4} \\ 2^{5-1}_{\text{V}} & \mathbf{I} = \pm\mathbf{1\ 2\ 3\ 4\ 5} \end{array} \tag{7}$$

In the above a *word* refers to a combination of elements such as **1 2 3**, **1 2 3 4**. In general the *resolution* of a design is equal to the smallest number of characters in any word appearing in the defining relation.

4: Resolution III Designs

Designs of resolution III are available which require only N runs to study up to $N - 1$ variables, where N is a multiple of four. We first discuss the arrangements for which N is a power of two. Particularly important designs are those for testing three variables in four runs, seven variables in eight runs and fifteen in sixteen runs. Two level designs for studying eleven variables in twelve runs, nineteen variables in twenty runs, etc., are derived by a somewhat different method due to Plackett & Burman (6), and are described later.

Designs for studying $k = N - 1$ variables in N runs may be called *saturated designs*. We introduce these designs by first considering a fractional for testing $k = 7$ variables in $N = 8$ runs. The complete factorial would require $2^7 = 128$ runs. We are considering therefore a one-sixteenth (i.e., a 2^{-4}) fractional, that is, a 2_{III}^{7-4} design. Since the design uses $2^3 = 8$ runs, we start construction of the design matrix with the 2^3 factorial, and then associate four additional variables with the plus and minus signs of the four interaction columns. For example, we may set

$$4 = 1\,2, \qquad 5 = 1\,3, \qquad 6 = 2\,3, \qquad 7 = 1\,2\,3 \tag{8}$$

to obtain the following 2_{III}^{7-4} design

TABLE 8
The Design Matrix for a 2_{III}^{7-4} Design

1	2	3	4 = 1 2	5 = 1 3	6 = 2 3	7 = 1 2 3
−	−	−	+	+	+	−
+	−	−	−	−	+	+
−	+	−	−	+	−	+
+	+	−	+	−	−	−
−	−	+	+	−	−	+
+	−	+	−	+	−	−
−	+	+	−	−	+	−
+	+	+	+	+	+	+

The identifications in Equation (8) provide the *generating relations*

$$\mathbf{I} = 1\,2\,4, \qquad \mathbf{I} = 1\,3\,5, \qquad \mathbf{I} = 2\,3\,6, \qquad \mathbf{I} = 1\,2\,3\,7 \tag{9}$$

associated with the generators **1 2 4**, **1 3 5**, **2 3 6** and **1 2 3 7**. Now clearly, if $\mathbf{I} = \mathbf{1\,2\,4}$ *and* $\mathbf{I} = \mathbf{1\,3\,5}$ then also $\mathbf{I} = \mathbf{1\,2\,4} \times \mathbf{1\,3\,5} = \mathbf{1}^2\,\mathbf{2\,3\,4\,5} = \mathbf{2\,3\,4\,5}$. Whence it follows for example that **2 3** and **4 5** are confounded. Thus, when there is more than one generator, the defining relation must contain not only the relations provided by the generators themselves, but all those obtained from all their possible products. The complete defining relation for this 2_{III}^{7-4} design is obtained, for example, by taking the generators first one at a time and then multiplying them together in all possible ways. Taking them one at a time gives $\mathbf{I} = \mathbf{1\,2\,4} = \mathbf{1\,3\,5} = \mathbf{2\,3\,6} = \mathbf{1\,2\,3\,7}$. Multiplying them together two at a time gives

$$\mathbf{I} = 2\,3\,4\,5 = 1\,3\,4\,6 = 3\,4\,7 = 1\,2\,5\,6 = 2\,5\,7 = 1\,6\,7,$$

three at a time gives:

$$\mathbf{I} = 4\,5\,6 = 1\,4\,5\,7 = 2\,4\,6\,7 = 3\,5\,6\,7,$$

and finally, four at a time gives:

$$\mathbf{I} = \mathbf{1\,2\,3\,4\,5\,6\,7}.$$

The complete *defining relation* for this 2^{7-4}_{III} design is therefore

$$\begin{aligned}\mathbf{I} &= \mathbf{1\,2\,4} = \mathbf{1\,3\,5} = \mathbf{2\,3\,6} = \mathbf{1\,2\,3\,7} = \mathbf{2\,3\,4\,5} = \mathbf{1\,3\,4\,6} = \mathbf{3\,4\,7}\\ &= \mathbf{1\,2\,5\,6} = \mathbf{2\,5\,7} = \mathbf{1\,6\,7} = \mathbf{4\,5\,6} = \mathbf{1\,4\,5\,7} = \mathbf{2\,4\,6\,7} = \mathbf{3\,5\,6\,7}\\ &= \mathbf{1\,2\,3\,4\,5\,6\,7}.\end{aligned} \tag{10}$$

As before, the defining relation quickly provides the alias structure for any effect, that is, indicates which effects are confounded. For example, multiplying the defining relation through by **1** we obtain

$$\begin{aligned}\mathbf{1} &= \mathbf{2\,4} = \mathbf{3\,5} = \mathbf{1\,2\,3\,6} = \mathbf{2\,3\,7} = \mathbf{1\,2\,3\,4\,5} = \mathbf{3\,4\,6} = \mathbf{1\,3\,4\,7} = \mathbf{2\,5\,6}\\ &= \mathbf{1\,2\,5\,7} = \mathbf{6\,7} = \mathbf{1\,4\,5\,6} = \mathbf{4\,5\,7} = \mathbf{1\,2\,4\,6\,7} = \mathbf{1\,3\,5\,6\,7} = \mathbf{2\,3\,4\,5\,6\,7}.\end{aligned}$$

Thus the interactions **2 4**, **3 5**, **1 2 3 6** etc., are seen to be aliases of, or confounded with, the main effect **1**. Similarly, multiplying through by **1 2 3** we obtain

$$\begin{aligned}\mathbf{1\,2\,3} &= \mathbf{3\,4} = \mathbf{2\,5} = \mathbf{1\,6} = \mathbf{7} = \mathbf{1\,4\,5} = \mathbf{2\,4\,6} = \mathbf{1\,2\,4\,7}\\ &= \mathbf{3\,5\,6} = \mathbf{1\,3\,5\,7} = \mathbf{2\,3\,6\,7} = \mathbf{1\,2\,3\,4\,5\,6} = \mathbf{2\,3\,4\,5\,7}\\ &= \mathbf{1\,3\,4\,6\,7} = \mathbf{1\,2\,5\,6\,7} = \mathbf{4\,5\,6\,7}.\end{aligned}$$

Thus the three-factor interaction **1 2 3** is an alias of, or confounded with **3 4**, **2 5**, **1 6**, etc. Since the resolution is determined by the smallest number of symbols forming any word in the defining relation, the design is of resolution III, as we have already noted.

In this example, if we write $\ell_1 = \frac{1}{4} \sum y\{\mathbf{1}\}$, $\ell_2 = \frac{1}{4} \sum y\{\mathbf{2}\}$, etc, and if we assume that all interactions between three of more variables are negligible, then by repeated use of the defining relation we obtain:

$$\begin{aligned}\ell_{\text{I}} &= \textbf{average}\\ \ell_1 &= \mathbf{1} + \mathbf{2\,4} + \mathbf{3\,5} + \mathbf{6\,7}\\ \ell_2 &= \mathbf{2} + \mathbf{1\,4} + \mathbf{3\,6} + \mathbf{5\,7}\\ \ell_3 &= \mathbf{3} + \mathbf{1\,5} + \mathbf{2\,6} + \mathbf{4\,7}\\ \ell_4 &= \mathbf{4} + \mathbf{1\,2} + \mathbf{5\,6} + \mathbf{3\,7}\\ \ell_5 &= \mathbf{5} + \mathbf{1\,3} + \mathbf{4\,6} + \mathbf{2\,7}\\ \ell_6 &= \mathbf{6} + \mathbf{2\,3} + \mathbf{4\,5} + \mathbf{1\,7}\\ \ell_7 &= \mathbf{7} + \mathbf{3\,4} + \mathbf{2\,5} + \mathbf{1\,6}.\end{aligned} \tag{11}$$

The Alternative Fractions

In writing down the design matrix for the 2^{7-4}_{III} fractional, the variables **4**, **5**, **6** and **7** were identified positively with the elements of the interactions **1 2**, **1 3**,

2 3 and **1 2 3** respectively. However, each of these identifications could have been made with either a plus or minus sign. For example, instead of associating the variable **4** positively with the interaction **1 2**, that is taking

$$+ \quad - \quad - \quad + \quad + \quad - \quad - \quad +$$

for its elements, the variable **4** could be associated negatively with the elements of **1 2**, that is:

$$- \quad + \quad + \quad - \quad - \quad + \quad + \quad -.$$

The first association gives $\mathbf{4} = \mathbf{1\,2}$ or equivalently $\mathbf{I} = \mathbf{1\,2\,4}$. The second association yields $\mathbf{4} = -\mathbf{1\,2}$ or equivalently $\mathbf{I} = -\mathbf{1\,2\,4}$. We could, in fact, have used any one of the sixteen identifications corresponding to the sixteen possible choices of signs

$$\mathbf{I} = \pm\mathbf{1\,2\,4}, \qquad \mathbf{I} = \pm\mathbf{1\,3\,5}, \qquad \mathbf{I} = \pm\mathbf{2\,3\,6}, \qquad \mathbf{I} = \pm\mathbf{1\,2\,3\,7}. \tag{12}$$

The sixteen possible identifications give the sixteen individual fractions which together yield the complete 2^7 design. In composing the defining relation for any one of the sixteen designs the usual rules of algebraic multiplication determine the signs in the defining relation and hence in the alias pattern.

Another one of these sixteen fractions is, for example, that in which variables **5** and **6** are associated with the elements of the interaction vectors **1 3** and **2 3** taken *negatively*, The generators for this design are:

$$\mathbf{1\,2\,4}, \qquad -\mathbf{1\,3\,5}, \qquad -\mathbf{2\,3\,6}, \qquad \mathbf{1\,2\,3\,7}, \tag{13}$$

and the corresponding defining relation is:

$$\begin{aligned} \mathbf{I} &= \mathbf{1\,2\,4} = -\mathbf{1\,3\,5} = -\mathbf{2\,3\,6} = \mathbf{1\,2\,3\,7} = -\mathbf{2\,3\,4\,5} = -\mathbf{1\,3\,4\,6} = \mathbf{3\,4\,7} \\ &= \mathbf{1\,2\,5\,6} = -\mathbf{2\,5\,7} = -\mathbf{1\,6\,7} = \mathbf{4\,5\,6} = -\mathbf{1\,4\,5\,7} = -\mathbf{2\,4\,6\,7} \\ &= \mathbf{3\,5\,6\,7} = \mathbf{1\,2\,3\,4\,5\,6\,7}. \end{aligned}$$

Assuming as before that all interactions between three or more variables are negligible, we see that this fraction allows the estimation of eight somewhat different combinations of effects

$$\begin{aligned} \ell'_{\mathrm{I}} &= \ \text{average} \\ \ell'_1 &= \ \mathbf{1} + \mathbf{2\,4} - \mathbf{3\,5} - \mathbf{6\,7} \\ \ell'_2 &= \ \mathbf{2} + \mathbf{1\,4} - \mathbf{3\,6} - \mathbf{5\,7} \\ \ell'_3 &= \ \mathbf{3} - \mathbf{1\,5} - \mathbf{2\,6} + \mathbf{4\,7} \\ \ell'_4 &= \ \mathbf{4} + \mathbf{1\,2} + \mathbf{5\,6} + \mathbf{3\,7} \\ \ell'_5 &= -\mathbf{5} + \mathbf{1\,3} - \mathbf{4\,6} + \mathbf{2\,7} \\ \ell'_6 &= -\mathbf{6} + \mathbf{2\,3} - \mathbf{4\,5} + \mathbf{1\,7} \\ \ell'_7 &= \ \mathbf{7} + \mathbf{3\,4} - \mathbf{2\,5} - \mathbf{1\,6} \end{aligned} \tag{14}$$

where the use of the prime notation on the ℓ's indicates only that some alternative function is under consideration. We see that Eq. (14) is identical to Eq. (11) with the numerals 5 and 6 having minus instead of plus signs.

Families of Fractionals

In the above example, there are $2^4 = 16$ different 2_{III}^{7-4} designs, each design corresponding to a particular choice of signs from among the generators $\pm\mathbf{1\ 2\ 4}$, $\pm\mathbf{1\ 3\ 5}$, $\pm\mathbf{2\ 3\ 6}$ and $\pm\mathbf{1\ 2\ 3\ 7}$. When the generators of a fractional factorial design associated with the identity $\mathbf{I}$ all have positive signs, they are called the *principal generators.* The defining relation obtained by multiplying out the generators is similarly called the *principal defining relation,* and the corresponding fractional factorial the *principal fraction.* Individual member fractions obtained from changes of sign in the generators are said to belong to the same *family.* In general, a 2^{k-p} fractional factorial design will have p generators, and the 2^p ways of allocating plus and minus signs to the generators will produce the 2^p different fractions belonging to the same family.

In general, a 2^{k-f} design will have f independent generators $\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_f$. An *independent* generator is such that it cannot be obtained by multiplying together the other generators, and is identified by the original association adopted in writing down the design. A defining relation for a particular fraction will contain 2^f words obtained by multiplying out $(\mathbf{I} \pm \mathbf{G}_1)(\mathbf{I} \pm \mathbf{G}_2) \cdots (\mathbf{I} \pm \mathbf{G}_f)$. The 2^f different fractions have defining relations given by the 2^f different ways of allocating plus and minus signs in this product. The defining relation for the *principal* fraction is given when all signs are plus. The alias pattern for any of the non-principal fractions is simply obtained by making the appropriate changes of sign in the alias pattern for the principal fraction.

Resolution III *Designs Containing* 16 *and* 32 *runs.*

The principal fraction of the 2_{III}^{15-11} design is obtained by first writing down the sixteen runs of the complete 2^4 design and then associating an additional eleven variables with the interactions **1 2**, **1 3**, **1 4**, **2 3**, **2 4**, **3 4**, **1 2 3**, **1 2 4**, **1 3 4**, **2 3 4**, and **1 2 3 4**. Similarly, the thirty-two runs comprising the 2_{III}^{31-26} factorial are obtained by writing down the complete factorial for five variables and then equating the additional twenty-six new variables with their interactions between the original five variables.

Effect of Dropping Variables

For intermediate values of k resolution III designs may be obtained by omitting variables from the resolution III design of next higher order. For example, to test six factors in eight runs we can use the 2_{III}^{7-4} design dropping out any one column in its design matrix. The alias relationships remain the same except that all words containing the characters associated with the dropped variables are omitted from the alias structure, and from any estimates of linear combinations. For example, dropping the columns **3** and **5** from the design matrix for the 2_{III}^{7-4} fractional given in Table 8 yields the 2_{III}^{5-2} design shown in Table 9. We can select the variables to be dropped out so that the most satisfactory alias arrangements exist among those remaining.

TABLE 9

Design Matrix 2^{5-2}_{III}

Defining Relation $\mathbf{I} = \mathbf{1\ 2\ 4} = \mathbf{1\ 6\ 7} = \mathbf{2\ 4\ 6\ 7}$.

1	2	4	6	7
−	−	+	+	−
+	−	−	+	+
−	+	−	−	+
+	+	+	−	−
−	−	+	−	+
+	−	−	−	−
−	+	−	+	−
+	+	+	+	+

Although it is true that a fractional of resolution R in a reduced number of $k - d$ variables can always be obtained by omitting d variables from a k variable fractional of resolution R, nevertheless a particular design obtained in this manner does not necessarily provide the best arrangement possible. For instance, if we drop variables **3**, **5**, **6** and **7** from the principal fraction 2^{7-4}_{III} design with generators **1 2 4**, **1 3 5**, **2 3 6** and **1 2 3 7** we are left with a design in the three variables **1**, **2** and **4** along with the unresolved generator **1 2 4** and hence the defining relation appropriate to a design having only *four* runs. On inspection we find that our eight factor combinations in the three remaining variables consist of *two replications* of the four run half-replicate design defined by $\mathbf{I} = \mathbf{1\ 2\ 4}$. This design is of resolution III, of course, but in many cases we would prefer to use the eight runs to perform a full factorial in the variables **1**, **2** and **4**. A full factorial would have been obtained had we, for example, dropped variables **1**, **2**, **3** and **7**.

The defining relation for the design obtained after dropping d variables will contain all those words in the original defining relation which do not contain any of the dropped numerals. Suppose among the f generators of the original design there are d generators that contain dropped variables, and $f - d$ generators that do not. A set of generators for the derived design will contain all the $f - d$ generators *not* containing dropped variables together with the largest set of independent products not containing dropped variables which can be found by multiplying the remaining generators.

For example consider again the resolution III design with generators $\mathbf{G}_1 = \mathbf{1\ 2\ 4}$, $\mathbf{G}_2 = \mathbf{1\ 3\ 5}$, $\mathbf{G}_3 = \mathbf{2\ 3\ 6}$ and $\mathbf{G}_4 = \mathbf{1\ 2\ 3\ 7}$. Suppose variable **1** is dropped. Since $\mathbf{G}_3 = \mathbf{2\ 3\ 6}$ does not contain **1** this generator will be included in the generators for the derived design. From the remaining generators we can obtain the products $\mathbf{G}_1\mathbf{G}_2$, $\mathbf{G}_1\mathbf{G}_4$ and $\mathbf{G}_2\mathbf{G}_4$ none of which contain the dropped variable **1**. Only two of the three products may be used since, having taken two of them, the third may be obtained by multiplication. For example, $\mathbf{G}_1\mathbf{G}_2 \cdot \mathbf{G}_1\mathbf{G}_4 = \mathbf{G}_2\mathbf{G}_4$. In general, a group of p words (such as the products we are considering here) are said to be independent if no one of them can be obtained by multiplying together some subset of the remaining $p - 1$. In this example then, a set of generators for the design derived after dropping **1** are **2 3 6**, **2 3 4 5** and **3 4 7** (that is, $\mathbf{G}_3$, $\mathbf{G}_1\mathbf{G}_2$ and $\mathbf{G}_1\mathbf{G}_4$).

At best, the effect of dropping d variables is to produce a design having d fewer generators. However, this represents the maximum reduction in generators possible, and particular choices of dropped variables may produce a smaller reduction in the number of generators. Of course, the greater the number of generators, the more words there will be in the defining relation and correspondingly, the more aliases for the remaining effects.

Effect of Combining Fractions from the Same Family

If we take the original fraction of the 2^{7-4}_{III} together with the second fraction in which the signs of **5** and **6** are switched, and take one-half the sums and differences of the respective linear combinations of effects we can estimate the following quantities (assuming all interactions with more than two factors to be nil).

From $\frac{1}{2}$ the Sums	From $\frac{1}{2}$ the Differences	
$\frac{1}{2}(\ell_{\text{I}} + \ell'_{\text{I}})$ = **Grand average**	$\frac{1}{2}(\ell_{\text{I}} - \ell'_{\text{I}})$ = **Block effect**	
$\frac{1}{2}(\ell_1 + \ell'_1) = \mathbf{1} + \mathbf{2\,4}$	$\frac{1}{2}(\ell_1 - \ell'_1) = \mathbf{3\,5} + \mathbf{6\,7}$	
$\frac{1}{2}(\ell_2 + \ell'_2) = \mathbf{2} + \mathbf{1\,4}$	$\frac{1}{2}(\ell_2 - \ell'_2) = \mathbf{3\,6} + \mathbf{5\,7}$	
$\frac{1}{2}(\ell_3 + \ell'_3) = \mathbf{3} + \mathbf{4\,7}$	$\frac{1}{2}(\ell_3 - \ell'_3) = \mathbf{1\,5} + \mathbf{2\,6}$	(15)
$\frac{1}{2}(\ell_4 + \ell'_4) = \mathbf{4} + \mathbf{1\,2} + \mathbf{5\,6} + \mathbf{3\,7}$	$\frac{1}{2}(\ell_4 - \ell'_4)$ = higher order interactions	
$\frac{1}{2}(\ell_5 + \ell'_5) = \mathbf{1\,3} + \mathbf{2\,7}$	$\frac{1}{2}(\ell_5 - \ell'_5) = \mathbf{5} + \mathbf{4\,6}$	
$\frac{1}{2}(\ell_6 + \ell'_6) = \mathbf{2\,3} + \mathbf{1\,7}$	$\frac{1}{2}(\ell_6 - \ell'_6) = \mathbf{6} + \mathbf{4\,5}$	
$\frac{1}{2}(\ell_7 + \ell'_7) = \mathbf{7} + \mathbf{3\,4}$	$\frac{1}{2}(\ell_7 - \ell'_7) = \mathbf{2\,5} + \mathbf{1\,6}$	

In general when two fractions from the same family are combined, the sums and differences of the corresponding linear combinations of the effects determine the effects which can be estimated from the combined design. The "block effect" referred to in Eq. (15) is the difference in average level between the first and second groups of eight runs.

Combining Fractional Factorials to Separate Effects

The procedure of adding fractions in sequence with suitably switched signs provides a useful method for the systematic isolation and confirmation of important effects in multi-variable systems. The method is very flexible and can be used in different ways as different situations unfold.

Mention will be made of two particular uses of this device: (1) the addition of a second fraction in which the signs in a *single column* are switched and, (2) the addition of a second fraction in which the signs in *all the columns* are switched.

Switching Signs for a Single Variable

Suppose a fractional factorial is generated by switching the signs associated with only the variable **1** in the 2^{7-4}_{III} factorial given in Table 8. Then the linear combinations that can be estimated from this fraction (given that the three-factor and higher order interactions are negligible) are the following

$$
\begin{aligned}
\ell_{\mathrm{I}} &= \text{Average} \\
\ell_1 &= -1 + 2\,4 + 3\,5 + 6\,7 \\
\ell_2 &= 2 - 1\,4 + 3\,6 + 5\,7 \\
\ell_3 &= 3 - 1\,5 + 2\,6 + 4\,7 \\
\ell_4 &= 4 - 1\,2 + 5\,6 + 3\,7 \\
\ell_5 &= 5 - 1\,3 + 4\,6 + 2\,7 \\
\ell_6 &= 6 + 2\,3 + 4\,5 - 1\,7 \\
\ell_7 &= 7 + 3\,4 + 2\,5 - 1\,6
\end{aligned}
\tag{16}
$$

Combining this fraction with the principal fraction, the following linear combination of effects are obtained from the combined design

From $\frac{1}{2}(\ell + \ell')$	From $\frac{1}{2}(\ell - \ell')$
Average	Block effect
2 4 + 3 5 + 6 7	1
2 + 3 6 + 5 7	1 4
3 + 2 6 + 4 7	1 5
4 + 5 6 + 3 7	1 2
5 + 4 6 + 2 7	1 3
6 + 2 3 + 4 5	1 7
7 + 3 4 + 2 5	1 6

(17)

We see that by adding to a fraction a further fraction with the signs for a single variable reversed, we isolate the main effect of that variable together with all of its two-factor interactions. Given any fractional of resolution III or higher and a second fractional identical to the first except that the signs of a single variable are switched, then the combined design will provide estimates of the main effect of the switched variable and all its associated two-factor interactions unbiased by any other main effect or two-factor interaction.

Switching Signs for All Variables

By switching signs for all seven variables given in the principal fraction we can estimate the following linear combinations

$$
\begin{aligned}
\ell'_{\mathrm{I}} &= \text{Average} \\
\ell'_1 &= -1 + 2\,4 + 3\,5 + 6\,7 \\
\ell'_2 &= -2 + 1\,4 + 3\,6 + 5\,7 \\
\ell'_3 &= -3 + 1\,5 + 2\,6 + 4\,7 \\
\ell'_4 &= -4 + 1\,2 + 5\,6 + 3\,7 \\
\ell'_5 &= -5 + 1\,3 + 4\,6 + 2\,7 \\
\ell'_6 &= -6 + 2\,3 + 4\,5 + 1\,7 \\
\ell'_7 &= -7 + 3\,4 + 2\,5 + 1\,6
\end{aligned}
\tag{18}
$$

By combining this fraction with the principal fraction all the main effects can be estimated clear of all the two-factor interactions. The two-factor interactions in turn will associate themselves in groups of three in accordance with the following scheme

From $\frac{1}{2}(\ell + \ell')$	From $\frac{1}{2}(\ell - \ell')$
Average	**Block effect**
2 4 + 3 5 + 6 7	1
1 4 + 3 6 + 5 7	2
1 5 + 2 6 + 4 7	3
1 2 + 5 6 + 3 7	4
1 3 + 4 6 + 2 7	5
2 3 + 4 5 + 1 7	6
3 4 + 2 5 + 1 6	7

(19)

This is a special example of a general principle, [14], which states that if any fractional is replicated with reversed signs, then all alias links between main effects and two-factor interactions are broken.

It should be noticed that although there are $2^7 = 128$ ways of switching signs, there are only $2^4 = 16$ of these switches that result in different designs. This must be so since there are only 2^4 different 2^{7-4} fractions belonging to the same family. It is easily confirmed by actual trial that the same design can be produced by a number of alternative sign switching arrangements, although the order in which the experimental runs appear may be different. The situation is made clear by considering only the generating relations for the principal fraction of the 2^{7-4}_{III}, that is:

$$\mathbf{I} = \mathbf{1\,2\,4}, \qquad \mathbf{I} = \mathbf{1\,3\,5}, \qquad \mathbf{I} = \mathbf{2\,3\,6}, \qquad \mathbf{I} = \mathbf{1\,2\,3\,7}.$$

It will be obvious for example that switching the signs of variables **4**, **5**, **7**, or of variable **1**, produces exactly the same effect. In each case the generating relations are:

$$\mathbf{I} = -\mathbf{1\,2\,4}, \qquad \mathbf{I} = -\mathbf{1\,3\,5}, \qquad \mathbf{I} = \mathbf{2\,3\,6}, \qquad \mathbf{I} = -\mathbf{1\,2\,3\,7}$$

Generators for Aggregate Designs

Suppose the principal fraction of the 2^{7-4}_{III} given in Table 8 is run. The generating relations for this design are

$$\mathbf{I}_8 = \mathbf{1\,2\,4}, \qquad \mathbf{I}_8 = \mathbf{1\,3\,5}, \qquad \mathbf{I}_8 = \mathbf{2\,3\,6}, \qquad \mathbf{I}_8 = \mathbf{1\,2\,3\,7}$$

where the notation $\mathbf{I}_8$ refers to a column of eight plus signs. Now suppose we perform a further series using a second 2^{7-4}_{III} from the same family as, for example, the fraction in which the variable **1** is run with reversed signs. The combined design formed from the two pieces is now a 2^{7-3} factorial. Since it is a one-eight replicate, it will have three generators, not four. How can these generators be identified? We note now that the generators for the second fraction are

$$\mathbf{I}_8 = -\mathbf{1\ 2\ 4}, \qquad \mathbf{I}_8 = -\mathbf{1\ 3\ 5}, \qquad \mathbf{I}_8 = \mathbf{2\ 3\ 6}, \qquad \mathbf{I}_8 = -\mathbf{1\ 2\ 3\ 7}.$$

It is clear that the generator **2 3 6** must be one of the generators for the combined design for in both pieces of the design $\mathbf{I}_8 = \mathbf{2\ 3\ 6}$. Consequently if $\mathbf{I}_{16}$ represents the column of sixteen plus signs associated with the complete design, then also $\mathbf{I}_{16} = \mathbf{2\ 3\ 6}$.

In asking what are the generating relations for the complete design we must first ask the question, "For which combinations are the products of the elements everywhere equal to $\mathbf{I}_{16}$?" Now we observe that **1 2 3 7** has the value $\mathbf{I}_8$ in the first set of eight runs, and $-\mathbf{I}_8$ in the second set. Thus, **1 2 3 7** is not equal to $\mathbf{I}_{16}$ and is therefore not a generator of the combined design. Similarly, **1 2 4** and **1 3 5** also are not generators for the complete design.

Now clearly for the first part of the design

$$\mathbf{I}_8 = (\mathbf{1\ 2\ 4})(\mathbf{1\ 3\ 5}) = \mathbf{2\ 3\ 4\ 5},$$

and also for the second part

$$\mathbf{I}_8 = (-\mathbf{1\ 2\ 4})(-\mathbf{1\ 3\ 5}) = \mathbf{2\ 3\ 4\ 5}.$$

Thus it is true for the complete design that

$$\mathbf{I}_{16} = \mathbf{2\ 3\ 4\ 5}.$$

Similarly, multiplying **1 2 4** by **1 2 3 7** it is true for the complete design that

$$\mathbf{I}_{16} = \mathbf{3\ 4\ 7}.$$

A third product is possible, obtained by multiplying **1 3 5** by **1 2 3 7** to give

$$\mathbf{I}_{16} = \mathbf{2\ 5\ 7}.$$

Now $(\mathbf{2\ 3\ 4\ 5})(\mathbf{3\ 4\ 7}) = \mathbf{2\ 5\ 7}$ and since it is a property of generators that no individual generator can be obtained from the others, we include in the new set of generators any two of the three derived above. Thus, the generating relations for this 2^{7-3} design are

$$\mathbf{I}_{16} = \mathbf{2\ 3\ 6}, \qquad \mathbf{I}_{16} = \mathbf{2\ 3\ 4\ 5}, \qquad \mathbf{I}_{16} = \mathbf{3\ 4\ 7}$$

and the corresponding defining relation is

$$\mathbf{I}_{16} = \mathbf{2\ 3\ 6} = \mathbf{2\ 3\ 4\ 5} = \mathbf{3\ 4\ 7} = \mathbf{4\ 5\ 6} = \mathbf{2\ 4\ 6\ 7} = \mathbf{2\ 5\ 7} = \mathbf{3\ 5\ 6\ 7}$$

From the above it will be seen that a general rule for finding generators for a design derived from two fractions from the same family each defined by generating relations of the kind

$$\text{Fraction 1: } I = \pm A = \pm B = \pm C = \cdots$$

$$\text{Fraction 2: } I = \pm A = \pm B = \pm C = \cdots$$

is as follows:

Suppose there are U words of unlike sign and L words of like sign in the two identities. Then $U + L - 1$ words which are generators of the new design will contain the L words of like sign together with $U - 1$ words obtained as independent even products of the U words of unlike sign.

In the above an *even* product is a product between an even number of words (usually two). This rule can be applied quite generally not only for combining designs of resolution III, but for combining any pair of fractionals belonging to the same family. As a further example, suppose two 2_{III}^{7-4} fractions were combined with generating relations:

$$\mathbf{I} = -\mathbf{1\,2\,4}, \qquad \mathbf{I} = -\mathbf{1\,3\,5}, \qquad \mathbf{I} = \mathbf{2\,3\,6}, \qquad \mathbf{I} = -\mathbf{1\,2\,3\,7}$$

and

$$\mathbf{I} = -\mathbf{1\,2\,4}, \qquad \mathbf{I} = \mathbf{1\,3\,5}, \qquad \mathbf{I} = -\mathbf{2\,3\,6}, \qquad \mathbf{I} = \mathbf{1\,2\,3\,7}$$

(The first fraction can be obtained from the principal fraction of the 2_{III}^{7-4} by reversing the sign of variable **1**, the second fraction by reversing the signs of variables **1** and **3**.) Then the generators for the complete design are $-\mathbf{1\,2\,4}$ and any two of the three words obtained from the even products of $-\mathbf{1\,3\,5}$, $\mathbf{2\,3\,6}$ and $-\mathbf{1\,2\,3\,7}$ to give the generating relations:

$$\mathbf{I} = -\mathbf{1\,2\,4}, \qquad \mathbf{I} = -\mathbf{1\,2\,5\,6}, \qquad \mathbf{I} = \mathbf{2\,5\,7}.$$

The reader will notice that switching to an alternative set of permissable generators leaves the design unchanged for it produces the same defining relation. Thus, in the above, if we had used the generators $-\mathbf{1\,2\,4}$, $-\mathbf{1\,2\,5\,6}$ and $-\mathbf{1\,6\,7}$ the defining relation obtained by multiplying out these generators would have been identical to that obtained before.

Alternative Choice of Generators

A particular fractional has an *unique* defining relation for a given design. There are however a number of different but equivalent choices of generators all of which lead to the same defining relation and the same design. Therefore, although we may speak of *the* defining relation for a design, we should properly refer to *a* choice of generators. In general, suppose $\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_f$ are a set of generators, necessarily independent, for a particular design. Then any other set of f independent generators derived by multiplication will be equivalent and will produce the same defining relation and be associated with the same design. The generators satisfy the same rules of multiplication as before, that is, $\mathbf{G}_1^2 = \mathbf{G}_2^2 = \cdots = \mathbf{G}_f^2 = \mathbf{I}$. To see that this is so suppose that $\mathbf{G}_1 = \mathbf{1\,2\,3}$, then $\mathbf{G}_1^2 = \mathbf{1}^2\mathbf{2}^2\mathbf{3}^2 = \mathbf{I}$. If we have four generators $\mathbf{G}_1$, $\mathbf{G}_2$, $\mathbf{G}_3$ and $\mathbf{G}_4$ for a particular design, then $\mathbf{G}_1\mathbf{G}_2$, $\mathbf{G}_1\mathbf{G}_3$, $\mathbf{G}_1\mathbf{G}_4$ and $\mathbf{G}_1\mathbf{G}_2\mathbf{G}_3$ will be an alternative set of generators, but $\mathbf{G}_1\mathbf{G}_2$, $\mathbf{G}_1\mathbf{G}_3$, $\mathbf{G}_1\mathbf{G}_4$, and $\mathbf{G}_1\mathbf{G}_2\mathbf{G}_3\mathbf{G}_4$ will not since $\mathbf{G}_1\mathbf{G}_2 \cdot \mathbf{G}_1\mathbf{G}_3 \cdot \mathbf{G}_1\mathbf{G}_4 = \mathbf{G}_1\mathbf{G}_2\mathbf{G}_3\mathbf{G}_4$. In particular, suppose we are interested in the fully saturated 2_{III}^{7-3} design with generators $\mathbf{G}_1 = \mathbf{1\,2\,4}$, $\mathbf{G}_2 = \mathbf{1\,3\,5}$, $\mathbf{G}_3 = \mathbf{2\,3\,6}$ and $\mathbf{G}_4 = \mathbf{1\,2\,3\,7}$, then the first legimate alternative set of generators will give $\mathbf{2\,3\,4\,5}$, $\mathbf{1\,3\,4\,6}$, $\mathbf{3\,4\,7}$ and $\mathbf{4\,5\,6}$ whereas the second "illigimate" choice gives $\mathbf{2\,3\,4\,5}$, $\mathbf{1\,3\,4\,6}$, $\mathbf{3\,4\,7}$ and $\mathbf{1\,2\,3\,4\,5\,6\,7}$, for it is readily confirmed that the last generator is the product of the first three.

Combining Fractionals Not of the Same Family

We have seen how by switching signs, fractional factorials may be combined together to isolate particular effects of interest, and that when fractional designs have the same generators except for their signs they are classified as being from the same family. Another method for isolating effects that is often of value is to combine fractions which are not of the same family. In one interesting species the numbers are switched in the generators as well as the signs. Possibilities arising from designs of this sort are presently being investigated.

Blocking Designs of Resolution III

Frequently an experimenter may fear that his results may be upset by shifts in average performance that occur from day to day, or with different batches of raw material. Such systematic sources of variation can often be successfully eliminated without biasing the estimates of the effects, or inflating the error variance by grouping the runs into "blocks".

The resolution III designs can be broken into two blocks of equal size by identifying the two blocks with the + and − versions of a single variable. For example, using the principal fraction of the 2_{III}^{7-4} design with generators **1 2 4**, **1 3 5**, **2 3 6** and **1 2 3 7** and using variable **7** for blocking we have the design given in Table 10. This design is a 2_{III}^{6-3} in blocks of four runs each. The generators

TABLE 10

1	2	3	4	5	6	7 = B	
+	−	−	−	−	+	+	Block 1
−	+	−	−	+	−	+	
−	−	+	+	−	−	+	
+	+	+	+	+	+	+	
−	−	−	+	+	+	−	Block 2
+	+	−	+	−	−	−	
+	−	+	−	+	−	−	
−	+	+	−	−	+	−	

for the design can now be written

$$\mathbf{1\,2\,4}, \qquad \mathbf{1\,3\,5}, \qquad \mathbf{2\,3\,6} \text{ and } \mathbf{1\,2\,3\,B} \tag{20}$$

where the letter **B** replaces the numeral **7** in the last generator to indicate the blocking variable. Assuming that three factor and higher order interactions are negligible, the defining relation for this design shows that the six main effects **1**, **2**, **3**, $\cdots$, **6** are each confounded with three two-factor interactions, one of which is a two-factor interaction with the blocks. The block effect itself is confounded with three two-factor interactions among the variables. In general, any resolution III design can be broken into two blocks of equal size by selecting the + and − signs of any one of the variables in the design matrix to identify the two blocks.

Resolution III designs can be broken into four blocks of equal size by identifying two block variables $\mathbf{B}_1$ and $\mathbf{B}_2$ with the + and − versions of two of the

TABLE 11

Run Number	Variables $1 = B_1 B_2$	2	3	4	5	Blocking Variables $6 = B_1$	$7 = B_2$		Block Variables B_1	B_2	$B_1 B_2$
1	+	+	−	+	−	−	−	Block 1	−	−	+
2	+	−	+	−	+	−	−				
3	−	−	−	+	+	+	−	Block 2	+	−	−
4	−	+	+	−	−	+	−				
5	−	+	−	−	+	−	+	Block 3	−	+	−
6	−	−	+	+	−	−	+				
7	+	−	−	−	−	+	+	Block 4	+	+	+
8	+	+	+	+	+	+	+				

variables. For example, starting with the principal fraction of the 2_{III}^{7-4} design and using variables **6** and **7** for blocking we obtain the four blocks of two runs each as illustrated in Table 11. Among these four blocks there are three degrees of freedom associated with the main effects and the two-factor interaction of pseudo-block variables $\mathbf{B}_1$, $\mathbf{B}_2$ identified with the two-way table

		$\mathbf{B}_1$ −	$\mathbf{B}_1$ +
$\mathbf{B}_2$	−	1, 2	3, 4
	+	5, 6	7, 8

The pairs of numbers in the cells of the table denote the runs comprising the four individual blocks. The "interaction variable" $\mathbf{B}_1 \mathbf{B}_2$ has precisely the same importance as the main effects $\mathbf{B}_1$ and $\mathbf{B}_2$. We see that on associating $\mathbf{B}_1$ with variable **6** and $\mathbf{B}_2$ with variable **7** we automatically associate a comparison between blocks, that is, the interaction $\mathbf{B}_1 \mathbf{B}_2$ with the interaction **6 7**. In this particular example **1** = **6 7** and hence the plus and minus signs of column **1** are now no longer available to accommodate an experimental variable. The variable **1**, therefore, is dropped from the experimental design. Thus, using variables **6**, **7** and **6 7** to identify the four blocks we obtain the design in the variables **2**, **3**, **4** and **5** in four blocks of two, as shown in Table 12.

It should be noted here that the two runs comprising each block are "mirror images" of one another, that is, within a block the versions of one run are exactly reversed in the second run. We will later see that this attribute of blocks of size two has important consequences.

It is usually assumed that block variables corresponding to such characteristics as the time of day, batches of raw material, operators, etc., do not interact with

TABLE 12

2	3	4	5	
+	−	+	−	
−	+	−	+	
−	−	+	+	The 2_{IV}^{4-1} in four blocks of two runs each
+	+	−	−	
+	−	−	+	
−	+	+	−	
−	−	−	−	
+	+	+	+	

the experimental variables. As always it is wise to regard this as a supposition to be tentatively entertained. (Unchecked assumptions are never safe in an applied subject.) We can be reminded of our supposition by setting out the analysis as if the supposition were *not* true, that is, by taking $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_1\mathbf{B}_2$ as if they were capable of interacting with the variables. If we treat $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_1\mathbf{B}_2$ on the same basis as the experimental variables, we have for the generators of the design given in Table 12:

$$\mathbf{B}_1\mathbf{B}_2\ \mathbf{2}\ \mathbf{4}, \qquad \mathbf{B}_1\mathbf{B}_2\ \mathbf{3}\ \mathbf{5}, \qquad \mathbf{B}_1\ \mathbf{2}\ \mathbf{3} \tag{21}$$

As mentioned above, the "interaction" $\mathbf{B}_1\mathbf{B}_2$ between the psuedo block factors $\mathbf{B}_1$ and $\mathbf{B}_2$ represents a contrast between the blocks which is on exactly the same footing as $\mathbf{B}_1$ or $\mathbf{B}_2$. (A mere relabeling of the blocks could change the "interaction" contrast to a main effect contrast $\mathbf{B}_1$.) Consequently, the combination $\mathbf{B}_1\mathbf{B}_2$ must be treated as a group having the same status as a single variable. In particular, a word such as $\mathbf{B}_1\mathbf{B}_2\ \mathbf{1}$ must count as a two-factor interaction (between variable **1** and one of the block contrasts) and *not* as a three-factor interaction.

The generators given in Equation 21 can be used to construct the defining relation for the design. Assuming all three-factor and higher order interactions negligible we obtain the linear combinations of effects given in Table 13.

The bracketed values in Table 13 indicate the two-factor interactions which could, if they existed, bias the various effects. Usually of course, all interactions with blocks can be safely supposed to be negligible. On this assumption the

TABLE 13

$$
\begin{aligned}
\ell_I &= \text{Average} \\
\ell_1 &= \mathrm{B_1B_2} + (2\,4 + 3\,5) \\
\ell_2 &= \quad 2 + (\mathrm{B_1B_2}4 + \mathrm{B_1}3 + \mathrm{B_2}5) \\
\ell_3 &= \quad 3 + (\mathrm{B_1B_2}5 + \mathrm{B_1}2 + \mathrm{B_2}4) \\
\ell_4 &= \quad 4 + (\mathrm{B_1B_2}2 + \mathrm{B_1}5 + \mathrm{B_2}3) \\
\ell_5 &= \quad 5 + (\mathrm{B_1B_2}3 + \mathrm{B_1}4 + \mathrm{B_2}2) \\
\ell_6 &= \quad \mathrm{B_1} + (2\,3 + 4\,5) \\
\ell_7 &= \quad \mathrm{B_2} + (3\,4 + 2\,5)
\end{aligned}
$$

Linear Combinations of Effects Provided by 2_{IV}^{4-1} in Four Blocks of Two Runs Each.

main effects of the variables **2**, **3**, **4** and **5** in this design are clear of two-factor interactions and the design given in Table 12 is in fact of resolution IV, that is, a 2_{IV}^{4-1} fractional in four blocks of two runs each.

The Plackett and Burman Designs

The methods given here allow us to construct resolution III designs suitable for exploring $k = 3$ variables in $N = 4$ runs, $k = 7$ in $N = 8$, $k = 15$ in $N = 16$ and $k = 31$ in $N = 32$ runs. It was pointed out by Plackett and Burman in 1946 [6] that two version designs which gave uncorrelated estimates of first order effects were available for exploring $k = N - 1$ variables in N runs where N was any multiple of four, and they presented the design matrices for these designs for 4 up to 100 (except for the isolated case of $N = 92$). When N is a power of two, the designs provided by Plackett and Burman are identical with one or the other of the families of resolution III designs derived by the methods given above. For the cases $N = 12$, 20, 24, 28 and 36 however, the Plackett and Burman designs allow useful gaps to be filled and are presented below.

The rows of plus and minus signs given in Table 14A are used to construct the design matrices for $N = 12$, 20, 24 and 36 while the design matrix for $N = 28$ is constructed from the nine rows shows in Table 14B.

TABLE 14A

$k = 11$	$N = 12$	++−+++−−−+−
$k = 19$	$N = 20$	++−−++−+−+−+−−−−++−
$k = 23$	$N = 24$	+++++−+−++−−++−−+−+−−−−
$k = 35$	$N = 36$	−+−+++−−−+++++−+++−−+−−−−+−+−++−−+−

TABLE 14B

$k = 27 \quad N = 28$

A	B	C
+ − + + + + − − −	− + − − − + − − +	+ + − + − + + − +
+ + − + + + − − −	− − + + − − + − −	− + + + + − + + −
− + + + + + − − −	+ − − − + − − + −	+ − + − + + − + +
− − − + − + + + +	− − + − + − − − +	+ − + + + − + − +
− − − + + − + + +	+ − − − − + + − −	+ + − − + + + + −
− − − − + + + + +	− + − + − − − + −	− + + + − + − + +
+ + + − − − + − +	− − + − − + − + −	+ − + + − + + + −
+ + + − − − + + −	+ − − + − − − − +	+ + − + + − − + +
+ + + − − − − + +	− + − − + − + − −	− + + − + + + − −

To construct the designs for $N = 12$, 20, 24 and 36 the plus and minus signs appearing in the appropriate row of Table 14A are first written down as a column. A second column is obtained from the first by moving down the elements of the first column once, and placing the last element in first position. This procedure is then repeated, moving down the second column one element to produce the third, and so on until all k columns are obtained. Finally a *row* of minus signs is added to complete the design. Thus, for the case of $k = 11$ variables

TABLE 14C

1	2	3	4	5	6	7	8	9	10	11
+	−	+	−	−	−	+	+	+	−	+
+	+	−	+	−	−	−	+	+	+	−
−	+	+	−	+	−	−	−	+	+	+
+	−	+	+	−	+	−	−	−	+	+
+	+	−	+	+	−	+	−	−	−	+
+	+	+	−	+	+	−	+	−	−	−
−	+	+	+	−	+	+	−	+	−	−
−	−	+	+	+	−	+	+	−	+	−
−	−	−	+	+	+	−	+	+	−	+
+	−	−	−	+	+	+	−	+	+	−
−	+	−	−	−	+	+	+	−	+	+
−	−	−	−	−	−	−	−	−	−	−

in $N = 12$ runs, the design of Table 14C is obtained. To construct the design for $k = 27$, $N = 28$ the three blocks, A, B and C illustrated in Table 14B are written down cyclically

$$\begin{array}{ccc} A & B & C \\ C & A & B \\ B & C & A \end{array}$$

and these twenty-seven rows followed by a row of minus signs.

An Example

In the start up of a new manufacturing unit considerable difficulty was experienced at the filtration stage. Other similar units operated satisfactorily at other sites but this particular new unit, although apparently similar in most major respects to the other units, gave a crude product which required very much longer filtration times. A meeting was called to discuss possible explanations and to consider ways of curing the trouble. The following variables were proposed as being possibly responsible.

(1) *The water supply*: The new plant used piped water from the local municipal reservoir. An alternative but somewhat limited supply of water was available from a local well. It was proposed that the effect of changing to the well water should be tried since it was argued that the well water corresponded more closely to the water used at other sites.

(2) *Raw Material*: The raw material used was manufactured on the site and it was suggested that this might be in some way deficient. It was proposed that raw material which had been satisfactorily used in the manufacturing of the product at another site should be shipped in and tested locally.

(3) *Temperature of Filtration*: This was not thought to be a critical factor over the range involved and no special attempt to control this temperature

had been made. However, the physical arrangement of the new process was such that filtration was accomplished at a somewhat lower temperature than had been experienced at other plants. By temporarily covering pipes and equipment, provision could be made to raise the temperature to the level experienced elsewhere.

(4) *Hold up Time*: Prior to filtration the product was held in a stirred tank. The average period of hold-up in the new plant was somewhat less than that used in the other plants but it could be easily increased.

(5) *Recycle*: The only major difference between production facilities at the other plants and the present one lay in the introduction of a recycle stage which slightly increased conversion of the reagents prior to precipitation and filtration. Arguments were advanced which accounted for the longer filtration time in terms of this recycle stage. Arrangements could be made to temporarily eliminate the recycle stage.

(6) *Rate of Addition of Caustic Soda*: Immediately prior to filtration a quantity of caustic soda liquor was added resulting in precipitation of the product. The addition rate was somewhat faster with the new plant but it was possible to produce slower rates of addition.

(7) *Type of Filter Cloth*: The filter cloths employed in this plant were very similar to those used at the other sites. However, they did come from more recently supplied batch and it was suggested that their performance should be compared with cloths from previously supplied batches which were still available.

In the following design the minus version corresponds to the usual operation for the new plant and the plus version to the change. Thus we have

	−	+
(1) water	town	well
(2) raw material	on site	other
(3) temperature of filtration	low	high
(4) hold up time	low	high
(5) recycle	included	omitted
(6) rate of addition NaOH	fast	slow
(7) filter cloth	new	old

The 2^{7-4}_{III} design with generators

$$\mathbf{I = 1\,2\,5}, \qquad \mathbf{I = 1\,3\,6}, \qquad \mathbf{I = 2\,3\,7}, \qquad \mathbf{I = 1\,2\,3\,4}$$

was chosen. This design is equivalent to the 2^{7-4}_{III} design considered earlier, but is obtained by a different association of variables, that is, $\mathbf{5 = 12}$, $\mathbf{6 = 13}$, $\mathbf{7 = 23}$ and $\mathbf{4 = 123}$. Eight experiments run in random order gave the filtration times listed below.

	1	2	3	4	5	6	7	Filtration Time
1	−	−	−	−	+	+	+	68.4
2	+	−	−	+	−	−	+	77.7
3	−	+	−	+	−	+	−	66.4
4	+	+	−	−	+	−	−	81.0
5	−	−	+	+	+	−	−	78.6
6	+	−	+	−	−	+	−	41.2
7	−	+	+	−	−	−	+	68.7
8	+	+	+	+	+	+	+	38.7

The usual analysis gives the estimates

water	$\ell_1 = \mathbf{1} + \mathbf{2\,5} + \mathbf{3\,6} + \mathbf{4\,7} = -10.9$
raw material	$\ell_2 = \mathbf{2} + \mathbf{1\,5} + \mathbf{3\,7} + \mathbf{4\,6} = -\ 2.8$
temperature	$\ell_3 = \mathbf{3} + \mathbf{1\,6} + \mathbf{2\,7} + \mathbf{4\,5} = -16.6$
hold up	$\ell_4 = \mathbf{4} + \mathbf{3\,5} + \mathbf{2\,6} + \mathbf{1\,7} = 0.5$
recycle	$\ell_5 = \mathbf{5} + \mathbf{1\,2} + \mathbf{3\,4} + \mathbf{6\,7} = 3.2$
rate of addition NaOH	$\ell_6 = \mathbf{6} + \mathbf{1\,3} + \mathbf{2\,4} + \mathbf{5\,7} = -22.8$
filter cloth	$\ell_7 = \mathbf{7} + \mathbf{2\,3} + \mathbf{1\,4} + \mathbf{5\,6} = -\ 3.4$.

The estimates -10.9, -16.6, and -22.8 are suspiciously large when compared to the others. The simplest interpretation of the results would be that the main effect of the factors **1**, **3** and **6** were important. However, many other interpretations are possible. Among these would be that the main effects of factor **3** and **6** and the interaction **3 6** (which is associated with **1**) were responsible for the observed results. Equivalently the main effects of **1** and **6** with **1 6**, or **1** and **3** with **1 3**, could be responsible. It was decided therefore to repeat the design with reverse signs, yielding the following results:

1	2	3	4	5	6	7	Filtration
+	+	+	+	−	−	−	66.7
−	+	+	−	+	+	−	65.0
+	−	+	−	+	−	+	86.4
−	−	+	+	−	+	+	61.9
+	+	−	−	−	+	+	47.8
−	+	−	+	+	−	+	59.0
+	−	−	+	+	+	−	42.6
−	−	−	−	−	−	−	67.6.

The estimates from this second design alone are

$$\begin{aligned}
\ell_1' &= -\mathbf{1} + \mathbf{2\,5} + \mathbf{3\,6} + \mathbf{4\,7} = 2.5\\
\ell_2' &= -\mathbf{2} + \mathbf{1\,5} + \mathbf{3\,7} + \mathbf{4\,6} = 5.0\\
\ell_3' &= -\mathbf{3} + \mathbf{1\,6} + \mathbf{2\,7} + \mathbf{4\,5} = -15.8\\
\ell_4' &= -\mathbf{4} + \mathbf{3\,5} + \mathbf{2\,6} + \mathbf{1\,7} = 9.2\\
\ell_5' &= -\mathbf{5} + \mathbf{1\,2} + \mathbf{3\,4} + \mathbf{6\,7} = -\ 2.3\\
\ell_6' &= -\mathbf{6} + \mathbf{1\,3} + \mathbf{2\,4} + \mathbf{5\,7} = 15.6\\
\ell_7' &= -\mathbf{7} + \mathbf{2\,3} + \mathbf{1\,4} + \mathbf{5\,6} = -\ 3.3
\end{aligned}$$

Whence by taking sums and differences of the linear combinations provided by the two component designs we obtain for the aggregate design

$$\begin{aligned}
\mathbf{1} &= -\ 6.7 & \mathbf{2\,5} + \mathbf{3\,6} + \mathbf{4\,7} &= -\ 4.2\\
\mathbf{2} &= -\ 3.9 & \mathbf{1\,5} + \mathbf{3\,7} + \mathbf{4\,6} &= 1.1\\
\mathbf{3} &= -\ 0.4 & \mathbf{1\,6} + \mathbf{2\,7} + \mathbf{4\,5} &= -16.2\\
\mathbf{4} &= -\ 4.4 & \mathbf{3\,5} + \mathbf{2\,6} + \mathbf{1\,7} &= 4.9\\
\mathbf{5} &= 2.8 & \mathbf{1\,2} + \mathbf{3\,4} + \mathbf{6\,7} &= 0.5\\
\mathbf{6} &= -19.2 & \mathbf{1\,3} + \mathbf{2\,4} + \mathbf{5\,7} &= -\ 3.6\\
\mathbf{7} &= 0.1 & \mathbf{2\,3} + \mathbf{1\,4} + \mathbf{5\,6} &= -\ 3.4
\end{aligned}$$

on review it seemed likely that the effect -19.2 associated with factor **6**, and the effect -16.2 associated with the linear combination (**1 6** + **2 7** + **4 5**) were probably real. It was also to be noted that the largest of the remaining effects, -6.7, was associated with factor **1**. The most likely explanation of the data

Rate of Addition of NaOH	Slow	65.4	42.6
	Fast	68.5	78.0
		Reservoir	Well
		Water Supply	

FIGURE 1
Two way table of average responses

therefore was that variables **1** and **6** both have effects and that they interacted. A two-way table of average values exemplifying these effects is shown below in Figure 1. It should be noted that the other explanations of the data are quite possible. For example the large effect attributed to the interaction between the factors **1** and **6** could be attributed equally well to the interaction **2 7** or **4 5**. The fact that none of the factors **2**, **4**, **5** and **7** have main effects does not, of course, preclude the possibility that their interactions exist. In fact in terms of the response surface if the center conditions of the experiment are located on the crest of a diagonally running ridge we should expect exactly this situation to occur. Of the possible explanations, however, that involving **1** and **6** and their two-factor interaction seemed by far the most likely. The crucial test was whether the trouble would be cured by using well water and the slowaddition rate of caustic soda while leaving the other variables at their usual levels.

A number of additional trials were run on the plant in which the only modifications made were the use of well water with a slow rate of addition of caustic. These runs did give satisfactorily short filtration times in the neighborhood of forty minutes and the modification was adopted.

5: Resolution IV Designs

We have seen that a valuable design can be generated by switching the signs of all the variables in a 2_{III}^{7-4} fractional factorial and adding the resultant design to the original fraction. This aggregate design, which uses sixteen runs, makes it possible to estimate all seven main effects clear of the two-factor interactions. The design is thus of resolution IV. In fact, it is a 2_{IV}^{7-3} design. It is possible to do even slightly better than this. The signs of the elements corresponding to the identity column I can also be switched, and the resulting set of eight positive and eight negative signs can be associated with an eighth factor. The final design is shown in Table 15 on page 338.

We call such a design a "fold over" design.

We must now consider what the generators and hence the defining relations are for this design. Each component group of eight runs can be regarded as a 2^{8-5} design with generating relations

$$\mathbf{I} = \mathbf{8} = \mathbf{1\,2\,4} = \mathbf{1\,3\,5} = \mathbf{2\,3\,6} = \mathbf{1\,2\,3\,7}$$

and

$$\mathbf{I} = -\mathbf{8} = -\mathbf{1\,2\,4} = -\mathbf{1\,3\,5} = -\mathbf{2\,3\,6} = \mathbf{1\,2\,3\,7}$$

TABLE 15
A 2^{8-4}_{IV} fold over design

8				4	5	6	7	
I	**1**	**2**	**3**	**1 2**	**1 3**	**2 3**	**1 2 3**	
+	−	−	−	+	+	+	−	
+	+	−	−	−	−	+	+	
+	−	+	−	−	+	−	+	
+	+	+	−	+	−	−	−	Principal fraction 2^{7-4}_{III}
+	−	−	+	+	−	−	+	
+	+	−	+	−	+	−	−	
+	−	+	+	−	−	+	−	
+	+	+	+	+	+	+	+	
−	+	+	+	−	−	−	+	
−	−	+	+	+	+	−	−	
−	+	−	+	+	−	+	−	
−	−	−	+	−	+	+	+	Principal fraction with all signs reversed
−	+	+	−	−	+	+	−	
−	−	+	−	+	−	+	+	
−	+	−	−	+	+	−	+	
−	−	−	−	−	−	−	−	

respectively. Applying the rule for combining fractions we notice at once that **1 2 3 7** is a generator for the aggregate design, and the remaining three generators are independent even products of **8**, **1 2 4**, **1 3 5**, and **2 3 6**. In particular, we can use **1 2 4 8**, **1 3 5 8** and **2 3 6 8** so that finally the generating relations for the aggregate design is:

$$\mathbf{I} = \mathbf{1\,2\,4\,8} = \mathbf{1\,3\,5\,8} = \mathbf{2\,3\,6\,8} = \mathbf{1\,2\,3\,7} \tag{22}$$

The defining relation is therefore

$$\begin{aligned}\mathbf{I} &= \mathbf{1\,2\,4\,8} = \mathbf{1\,3\,5\,8} = \mathbf{2\,3\,6\,8} = \mathbf{1\,2\,3\,7} = \mathbf{2\,3\,4\,5} = \mathbf{1\,3\,4\,6} = \mathbf{3\,4\,7\,8}\\ &= \mathbf{1\,2\,5\,6} = \mathbf{2\,5\,7\,8} = \mathbf{1\,6\,7\,8} = \mathbf{4\,5\,6\,8} = \mathbf{2\,4\,6\,7} = \mathbf{1\,4\,5\,7} = \mathbf{3\,5\,6\,7}\\ &= \mathbf{1\,2\,3\,4\,5\,6\,7\,8}\end{aligned}$$

The generating relations for all sixteen of the 2^{8-4}_{IV} factionals are:

$$\mathbf{I} = \pm\mathbf{1\,2\,4\,8}, \quad \mathbf{I} = \pm\mathbf{1\,3\,5\,8}, \quad \mathbf{I} = \pm\mathbf{2\,3\,6\,8} \text{ and } \mathbf{I} = \pm\mathbf{1\,2\,3\,7}\,.$$

Ignoring interactions between three or more factors, and using the principal defining relation, the sixteen quantities which can now be estimated from the principal one-sixteenth fraction are given in Table 16.

As before further fractions can be performed in combination with the original fraction to isolate particular two-factor interactions or combinations of two-factor interactions. It will be seen now that when a design is formed containing 2^{k+1} runs from a design containing 2^k runs by replicating the 2^k design with reversed signs and associating some further factor **X** with the 2^k plus ones and 2^k minus ones, then a general rule for obtaining the generators and defining relation of the new design from the generators and defining relation of the old design is as follows: 1) All generators which contain an even number of

TABLE 16

Effects estimable using the 2_{IV}^{8-4} design

	Average
8 main effects	**1**
	2
	3
	4
	5
	6
	7
	8
7 sets of two-factor interactions confounded in groups of four	**1 2 + 3 7 + 4 8 + 5 6**
	1 3 + 2 7 + 5 8 + 4 6
	1 4 + 2 8 + 3 6 + 5 7
	1 5 + 3 8 + 2 6 + 4 7
	1 6 + 7 8 + 3 4 + 2 5
	1 7 + 2 3 + 6 8 + 4 5
	1 8 + 2 4 + 3 5 + 6 7

characters in the original design are retained as generators in the new design, 2) All generators which contain an odd number of characters in the original designs will be reproduced containing the extra character **X** as generators in the new design. For example, the generator **1 3 4** will become **1 3 4 X**.

An Alternative Method for Generating Designs of Resolution IV

An inspection of the generators for the 2_{IV}^{8-4} design just described will show that an alternative method for constructing this design would be to write down in standard order the sixteen combinations of variables for a complete 2^4 factorial, and then to associate further factors with the four three-factor interactions. To demonstrate, let the 2^4 factorial be written down in terms of the variables **1**, **2**, **3** and **8**. The four three-factor interactions are then **1 2 8**, **1 3 8**, **2 3 8** and **1 2 3**. These can now be associated with the four new variables **4**, **5**, **6** and **7** to give the set of four generators

$$\begin{array}{ccccc} \mathbf{1} & \mathbf{2} & & \mathbf{8} & \mathbf{4} \\ \mathbf{1} & & \mathbf{3} & \mathbf{8} & \mathbf{5} \\ & \mathbf{2} & \mathbf{3} & \mathbf{8} & \mathbf{6} \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & & \mathbf{7} \end{array} \tag{24}$$

The design thus constructed is identical to that given in Table 15. The only reason, of course, for starting off with variables **1**, **2**, **3** and **8** instead of **1**, **2**, **3** and **4** is to show the identity between this method of construction and the previous one.

As a further example of this second method for constructing resolution IV designs let us construct the 2_{IV}^{16-11} design. Since the design contains 32 runs we begin by writing down the full 2^5 factorial in the variables **1**, **2**, **3**, **4** and **5**. Eleven additional variables are now introduced by associating them with the ten three-factor interactions and the single five-factor interaction. We thus have for the set of eleven generators

$$\begin{array}{cccccc}
1 & 2 & 3 & & & 6 \\
1 & 2 & & 4 & & 7 \\
1 & 2 & & & 5 & 8 \\
1 & & 3 & 4 & & 9 \\
1 & & 3 & & 5 & 10 \\
1 & & & 4 & 5 & 11 \\
 & 2 & 3 & 4 & & 12 \\
 & 2 & 3 & & 5 & 13 \\
 & 2 & & 4 & 5 & 14 \\
 & & 3 & 4 & 5 & 15 \\
1 & 2 & 3 & 4 & 5 & 16
\end{array} \tag{25}$$

If three-factor and higher order interaction terms are negligible, thirty-two independent estimates can be obtained. They include the grand average, the sixteen main effects **1**, **2**, **3**, $\cdots$, **16**; and the fifteen combinations of two-factor interactions displayed below

1 2 + 15 16 + 3 6 + 4 7 + 5 8 + 9 12 + 10 13 + 11 14

1 3 + 2 6 + 14 16 + 4 9 + 5 10 + 11 15 + 7 12 + 8 13

1 4 + 2 7 + 3 9 + 13 16 + 5 11 + 6 12 + 10 15 + 8 14

1 5 + 2 8 + 3 10 + 4 11 + 12 16 + 6 13 + 7 14 + 9 15

1 6 + 2 3 + 14 15 + 4 12 + 5 13 + 11 16 + 7 9 + 8 10

1 7 + 2 4 + 3 12 + 13 15 + 5 14 + 6 9 + 10 16 + 8 11

1 8 + 2 5 + 3 13 + 4 14 + 12 15 + 6 10 + 7 11 + 9 16

1 9 + 2 12 + 3 4 + 10 11 + 5 15 + 6 7 + 13 14 + 8 16

1 10 + 2 13 + 3 5 + 4 15 + 12 14 + 6 8 + 7 16 + 9 11

1 11 + 2 14 + 3 15 + 4 5 + 9 10 + 6 16 + 7 8 + 12 13

1 12 + 2 9 + 3 7 + 4 6 + 5 16 + 10 14 + 8 15 + 11 13

1 13 + 2 10 + 3 8 + 4 16 + 5 6 + 11 12 + 7 15 + 9 14

1 14 + 2 11 + 3 16 + 4 8 + 5 7 + 6 15 + 9 13 + 10 12

1 15 + 2 16 + 3 11 + 4 10 + 5 9 + 6 14 + 7 13 + 8 12

1 16 + 2 15 + 3 14 + 4 13 + 5 12 + 6 11 + 7 10 + 8 9

In general, a resolution IV design may always be constructed by first writing down the design matrix for a two-level factorial and then associating new variables with all those interaction columns having an *odd* number of numerals.

Of course, this 2_{IV}^{16-11} design could have been obtained by fold-over by first writing down the 2_{III}^{15-11} design, the saturated resolution III design for fifteen variables in sixteen runs. The eleven generators for this design is given in Table 17a. In Table 17a the variables are numbered from 2 to 16 to make the equivalence between the two methods of construction evident. The generators for the 2_{IV}^{16-11} obtained by fold-over is shown in Table 17b. These generators are obtained by attaching the variable 1 to every word in the generating relation of the 2_{III}^{15-11} having an odd number of symbols. The generators, and hence the design obtained by fold-over, are thus identical to those displayed earlier in Eq(25). The same principle of fold-over may be used with the Plackett and Burman designs. For example, using the Plackett and Burman design for $k = 11$

TABLE 17a

Generating Relation 2_{III}^{15-11}

2	3			6
2		4		7
2			5	8
	3	4		9
	3		5	10
		4	5	11
2	3	4		12
2	3		5	13
2		4	5	14
	3	4	5	15
2	3	4	5	16

TABLE 17b

Generation Relation for 2_{IV}^{16-11} *Obtained by Fold-over*

1	2	3			6
1	2		4		7
1	2			5	8
1		3	4		9
1		3		5	10
1			4	5	11
	2	3	4		12
	2	3		5	13
	2		4	5	14
		3	4	5	15
1	2	3	4	5	16

variables in twelve runs we may derive a design usable for studying twelve variables in twenty-four runs in which no two-factor interaction is aliased with any main effect.

Complete Factorials within Fractionals Applied to Screening

When little is known about the variables which effect a particular response we are in what may be called a screening situation. That is to say, that although it is necessary to test a rather large number of variables which might conceivably have important effects, it can be realistically postulated that only a few, perhaps one, two or three of the variables, will be of major importance. Whichever variables do turn out to be of major influence may of course interact with one another. To put this argument in another way, we may have a fairly large number, say eight variables, which are of possible importance, but we believe

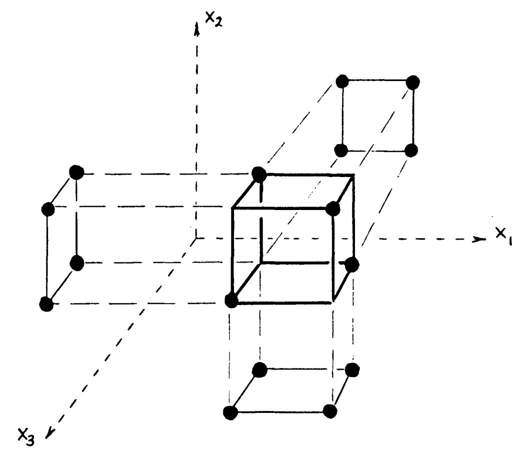

FIGURE 2—Projection of 2^{3-1}_{III} into three 2^2 factorials.

the effects of all but, say, three of these are likely to be negligible. Thus, we tentatively entertain the idea that at least five of the variables can be regarded essentially as dummies, but we don't know *which* five. In these circumstances we need a design in the complete set of eight variables which will produce a complete factorial in *any* three of the component variables. Thus, although we don't know which subset of the variables will turn out to be important, whichever subset does, provides a full factorial, or even a replicated factorial, in those variables.

The basic idea is illustrated in the very simpliest case for the one-half replicate of the 2^3 factorial shown in Figure 2. Suppose the total number of variables considered is three, but it can be reasonably postulated that not more than two have any real effects. Then we see from Figure 2 that the design supplies a complete factorial in any of the three pairs of variables since each projection of the 2^{3-1}_{III} design into a two dimensional plane produces a complete factorial design. This is also apparent from inspection of the design matrix since, if we drop any one column of the design matrix, the remaining two columns provide a full 2^2 factorial. This can be seen even more simply, for the generating relation for this design is $\mathbf{I} = \mathbf{1\ 2\ 3}$ and if any one of the variables is dropped the generator will vanish showing that the resulting design is not a fractional factorial.

In general, it is clear that a design of resolution R will provide a complete factorial in any sub-set of the $(R - 1)$ variables. This must be so since every word in the defining relation contains R or more characters. It follows that is all but $(R - 1)$ characters are treated as dummies, then every word in the defining relation will disappear.

2^3 *Factorials within Resolution* IV *Designs*

If a design of resolution IV contains $r \times 2^3$ runs then it can be regarded as providing r replicates of a full factorial in any three variables. As an example, consider the sixteen-run resolution IV design for eight variables i.e. the 2^{8-4}_{IV} design.

This design can be regarded as providing a twice replicated 2^3 factorial for every one of the fifty-six choices of three variables out of eight. Geometrically, this means that the sixteen points in eight dimensional space can be projected into any one of the fifty-six three dimensional coordinate sub-spaces to produce a replicated cube. The reader can readily confirm for himself that the omission of any five columns from Table 15 provides a twice replicated factorial in the remaining variables.

As always, evidence from experiments of this kind should only be regarded as suggestive and subject to confirmation rather than as supplying definite proof. Alternative explanations of the results obtained from such experiments involving higher order interactions could be easily produced. However, in selecting alternative explanations as worthy of further study we rest heavily upon our prior beliefs about the plausibility of these alternatives.

It is interesting to note an early use of designs of this kind by Tippett [15]. An adequate statement of the proper attitude towards the results is to be found in a discussion by R. A. Fisher [12] of Tippett's example.

General Rules for Designs Obtained by Projection

We have seen that a design of resolution R provides a complete factorial in any sub-set of $(R - 1)$ variables. In particular, designs of resolution III may be used for screening up to two variables out of $N - 1$ variables, designs of resolution IV may be used for screening up to three variables out of $N/2$ variables, and designs of resolution V, which we shall discuss later, may be used for screening up to four variables out of a larger number. If a design of resolution R is used to screen subsets of R variables, then full factorials will result for certain subsets, and fractional factorial for others. Those subsets of variables providing fractional factorials are simply subsets which appear as words in the final defining relation. For example, consider the 2_{IV}^{8-4} design discussed earlier. Its defining relation is

$$\begin{aligned} \mathbf{I} &= \mathbf{1248} = \mathbf{1358} = \mathbf{2368} = \mathbf{1237} = \mathbf{2345} = \mathbf{1346} = \mathbf{3478} \\ &= \mathbf{1256} = \mathbf{2578} = \mathbf{1678} = \mathbf{4568} = \mathbf{2467} = \mathbf{1457} = \mathbf{3567} \\ &= \mathbf{12345678}. \end{aligned}$$

Regarded as a design to screen sub-sets of four variables, this design will provide replicated half-fractions for the fourteen combinations of variables **1**, **2**, **4** and **8**; **1**, **3**, **5** and **8**; **2**, **3**, **6** and **8**; etc., which appear as forming words in the defining relation, and complete 2^4 factorials designs for any one of the remaining fifty-six combinations of four variables. In the case of resolution V designs we can, in accordance to our general rule, obtain full factorials in any set of four variables. These designs would, for most purposes, also be adequate for screening five variables because even for those combinations of variables which appear as words in the defining relation, one-half replicates would be available, and these would permit all main effects and two factor interactions to be distinguished, on the assumption of course that higher order interaction effects are negligible.

Example

The problem of analyzing these designs can be thought of either as picking out the one, two or three variables whose main effects and interactions can account for all the effects found, or equivalently for looking for sets of replicates within the runs. As an example, consider the data given in Table 18 obtained from a screening experiment containing eight variables, using the generators **1 2 4 8**, **1 3 5 8**, **2 3 6 8** and **1 2 3 7**

TABLE 18

Run #	1	2	3	8	4	5	6	7	
1	−	−	−	−	−	−	−	−	60.4
2	+	−	−	−	+	+	−	+	66.0
3	−	+	−	−	+	−	+	+	62.1
4	+	+	−	−	−	+	+	−	63.3
5	−	−	+	−	−	+	+	+	82.9
6	+	−	+	−	+	−	+	−	75.4
7	−	+	+	−	+	+	−	−	82.4
8	+	+	+	−	−	−	−	+	73.0
9	−	−	−	+	+	+	+	−	68.1
10	+	−	−	+	−	−	+	+	61.2
11	−	+	−	+	−	+	−	+	71.3
12	+	+	−	+	+	−	−	−	59.6
13	−	−	+	+	+	−	−	+	67.3
14	+	−	+	+	−	+	−	−	75.3
15	−	+	+	+	−	−	+	−	66.7
16	+	+	+	+	+	+	+	+	77.1

The estimated effects are given in Table 19a.

TABLE 19a

Average	69.5
1	−1.3
2	−0.1
3	11.0
4	0.5
5	7.6
6	0.2
7	1.2
8	−2.4
12 + 37 + 48 + 56	−1.1
13 + 27 + 58 + 46	1.7
14 + 28 + 36 + 57	0.8
15 + 38 + 26 + 47	−4.5
16 + 78 + 34 + 25	0.6
17 + 23 + 68 + 45	−0.3
18 + 24 + 35 + 67	1.2

TABLE 19b

3	8	5	Responses	
−	−	−	60.4	62.1
+	−	−	75.4	73.0
−	+	−	61.2	59.6
+	+	−	67.3	66.7
−	−	+	66.0	63.3
+	−	+	82.9	82.4
−	+	+	68.1	71.3
+	+	+	75.3	77.1

It was not expected that more than a few of the eight variables would in fact have important effects upon the response, and it will be seen that the data is readily explained by supposing that the important variables are **3**, **5** and **8**. The main effects and two-factor interactions associated with these variables are underlined in Table 19a. On this explanation runs 1 and 3, 2 and 4, 5 and 7, 6 and 8, 9 and 11, and 10 and 12, 13 and 15 and finally 14 and 16 are essentially duplicates one of the other differing mainly because of experimental error and partly because of effects of the other variables of lesser importance. The data are rearranged as a duplicated 2^3 factorial in variables **3**, **8** and **5** in Table 19b.

In an experiment of this kind it would have been advantageous to have available some independent estimate of pure error obtained, for example, from duplication of certain of the runs selected in accordance with principles described elsewhere [10]. In such a case we could then compare the size of the error obtained from the "constructed" duplicates with that from known duplicates.

Blocking for Designs of Resolution IV

Assuming that interactions between three or more variables are negligible, the 2_{IV}^{8-4} design with generators **1 2 3 7**, **1 2 4 8**, **1 3 5 8** and **2 3 6 8** provides independent estimates of the eight main effects and of seven groups of two-factor interactions. By using the $+$ and $-$ signs associated with the interaction columns, this design can be broken into either two, four or eight equal sized blocks which are unconfounded with main effects.

For example, we may use the $+$ and $-$ signs associated with the two-factor interaction set $\mathbf{1\,2 + 3\,7 + 4\,8 + 5\,6}$, to define two blocks, if we call the block contrast $\mathbf{B}_1$, and put $\mathbf{B}_1 = \mathbf{1\,2 + 3\,7 + 4\,8 + 5\,6}$. To break the design into four equal blocks, two columns associated with the interaction sets may be used. For example, we might choose $\mathbf{B}_1 = \mathbf{1\,2 + 3\,7 + 4\,8 + 5\,6}$ and $\mathbf{B}_2 = \mathbf{1\,3 + 2\,7 + 5\,8 + 4\,6}$. Each of the four blocks will contain the four runs identified by the pairs of versions $(\pm, \pm)$, that is, the four sets of versions $(-, -)$, $(+, -)$, $(-, +)$, $(+, +)$ provided by the two interaction columns. The block interaction effect $\mathbf{B}_1\mathbf{B}_2$ will then be found to be associated with another two-factor interaction set, that is, $\mathbf{B}_1\mathbf{B}_2 = \mathbf{1\,7 + 2\,3 + 6\,8 + 4\,5}$. This can be confirmed by actually multiplying out the elements of the columns associated with $\mathbf{B}_1$ and $\mathbf{B}_2$, or simply by noting, for example, that products of the interaction elements in $\mathbf{B}_1$ and $\mathbf{B}_2$ are $\mathbf{1\,2 \times 1\,3 = 2\,3}$, $\mathbf{1\,2 \times 2\,7 = 1\,7}$; $\mathbf{5\,6 \times 4\,8 = 4\,5}$ etc.

To break the design into eight equal blocks, three of the interaction sets must be used. However, in choosing the third set we may not use the column associated with the interaction $\mathbf{B}_1\mathbf{B}_2$, although any of the remaining interaction columns may be used. Let us choose $\mathbf{B}_3 = \mathbf{1\,8 + 2\,4 + 3\,5 + 6\,7}$. Each of the blocks will now contain the two runs identified by the eight sets of versions $(\pm, \pm, \pm)$ provided by $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_3$. It can be readily confirmed by multiplying out the elements of the block columns that the complete set of seven two-factor interaction comparisons are now used up, that is,

TABLE 20a

Construction of 2_{IV}^{8-4} in Blocks of Two

1	2	3	8	7 1 2 3	4 1 2 8	5 1 3 8	6 2 3 8	B_1 1 2	B_2 1 3	B_3 1 8
−	−	−	−	−	−	−	−	+	+	+
+	−	−	−	+	+	+	−	−	−	−
−	+	−	−	+	+	−	+	−	+	+
+	+	−	−	−	−	+	+	+	−	−
−	−	+	−	+	−	+	+	+	−	+
+	−	+	−	−	+	−	+	−	+	−
−	+	+	−	−	+	+	−	−	−	+
+	+	+	−	+	−	−	−	+	+	−
−	−	−	+	−	+	+	+	+	+	−
+	−	−	+	+	−	−	+	−	−	+
−	+	−	+	+	−	+	−	−	+	−
+	+	−	+	−	+	−	−	+	−	+
−	−	+	+	+	+	−	−	+	−	−
+	−	+	+	−	−	+	−	−	+	+
−	+	+	+	−	−	−	+	−	−	−
+	+	+	+	+	+	+	+	+	+	+

TABLE 20b

Construction of 2_{IV}^{8-4} in Blocks of Two

1	2	3	8	7	4	5	6	B_1	B_2	B_3	
+	−	−	−	+	+	+	−	−	−	−	Block 1
−	+	+	+	−	−	−	+				
+	+	−	−	−	−	+	+	+	−	−	Block 2
−	−	+	+	+	+	−	−				
+	−	+	−	−	+	−	+	−	+	−	Block 3
−	+	−	+	+	−	+	−				
+	+	+	−	+	−	−	−	+	+	−	Block 4
−	−	−	+	−	+	+	+				
+	−	−	+	+	−	−	+	−	−	+	Block 5
−	+	+	−	−	+	+	−				
+	+	−	+	−	+	−	−	+	−	+	Block 6
−	−	+	−	+	−	+	+				
+	−	+	+	−	−	+	−	−	+	+	Block 7
−	+	−	−	+	+	−	+				
+	+	+	+	+	+	+	+	+	+	+	Block 8
−	−	−	−	−	−	−	−				

$$
\begin{aligned}
\mathbf{B}_1 &= \mathbf{1\,2} + \mathbf{3\,7} + \mathbf{4\,8} + \mathbf{5\,6} \\
\mathbf{B}_2 &= \mathbf{1\,3} + \mathbf{2\,7} + \mathbf{5\,8} + \mathbf{4\,6} \\
\mathbf{B}_3 &= \mathbf{1\,8} + \mathbf{2\,4} + \mathbf{3\,5} + \mathbf{6\,7} \\
\mathbf{B}_1\mathbf{B}_2 &= \mathbf{1\,7} + \mathbf{2\,3} + \mathbf{6\,8} + \mathbf{4\,5} \\
\mathbf{B}_1\mathbf{B}_3 &= \mathbf{1\,4} + \mathbf{2\,8} + \mathbf{3\,6} + \mathbf{5\,7} \\
\mathbf{B}_2\mathbf{B}_3 &= \mathbf{1\,5} + \mathbf{3\,8} + \mathbf{2\,6} + \mathbf{4\,7} \\
\mathbf{B}_1\mathbf{B}_2\mathbf{B}_3 &= \mathbf{1\,6} + \mathbf{7\,8} + \mathbf{3\,4} + \mathbf{2\,5}
\end{aligned}
\tag{26}
$$

As the reader can confirm for himself, subject only to the condition that the $\mathbf{B}_3$ must not be chosen so as to coincide with $\mathbf{B}_1\mathbf{B}_2$, the association between blocks and two-factor interactions can be made in any other way whatever.

Tables 20a and 20b show how we can write out a 2_{IV}^{8-4} design arranged in eight blocks of two runs each. Since the complete design contains sixteen runs, we begin by writing down a 2^4 factorial in four of the eight variables. (In order that the final design may be compared with designs obtained previously, we chose these variables to be **1**, **2**, **3** and **8** although they could just as easily been chosen to be **1**, **2**, **3** and **4**.) The variables **4**, **5**, **6** and **7** are then associated with the three-factor interactions. The generators are **1 2 3 7**, **1 2 4 8**, **1 3 5 8** and **2 3 6 8**. As illustrated in Table 20a we now write down the three columns corresponding to the interactions **1 2**, **1 3**, and **1 8** and associate these with the block factors $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_3$. The eight blocks are then obtained by putting those pairs of runs for which $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_3$ are $(- \; - \; -)$ into the first block, the pair of runs for which $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_3$ are $(+ \; - \; -)$ in the second block and so on.

In Table 20b it will be noted that the second run in each block is the foldover, or mirror image, of the first run, that is, the versions of one run are exactly reversed in the second. Suppose now that a single run is taken from each block such that one of the variables always appears with the same sign. Choosing, for example, those runs with the + version of variable 1 we obtain the array given in Table 21.

The reader will note that the result, omitting variable **1**, is the 2_{III}^{7-4} design in the variables **2**, **3**, **4**, $\cdots$, **8** with generating relation $\mathbf{I} = \mathbf{2\,3\,7} = \mathbf{2\,8\,4} = \mathbf{3\,8\,5} = \mathbf{2\,3\,8\,6}$. This design is identical, except for the number identification given the seven variables, to the 2_{III}^{7-4} design described earlier.

TABLE 21

1	2	3	8	7	4	5	6
+	−	−	−	+	+	+	−
+	+	−	−	−	−	+	+
+	−	+	−	−	+	−	+
+	+	+	−	+	−	−	−
+	−	−	+	+	−	−	+
+	+	−	+	−	+	−	−
+	−	+	+	−	−	+	−
+	+	+	+	+	+	+	+

We see now that the principle of fold-over can be modified slightly to provide resolution IV designs automatically broken into blocks of two runs such that the block effects are unconfounded with the main effects. We begin by writing down the design matrix for the appropriate resolution III design in k variables plus an additional column **I** consisting solely of plus signs. Each row of the design matrix is then folded-over, that is, repeated with all signs reversed. The pairs of rows form blocks of two of a resolution IV design in $k + 1$ variables. For example, the 2_{IV}^{4-2} in blocks of two is constructed by first writing down the 2_{III}^{3-1} along with the column **I**, and then folding over each row to provide four blocks of two runs each as illustrated in Table 22. This 2_{IV}^{4-1} design is identical

TABLE 22

Original 2^{3-1}, I = 1 2 3				2_{IV}^{4-1} in Blocks of two, I = 1 2 3 4				
1	2	3	I	1	2	3	4	
−	−	+	+	−	−	+	+	Block 1
+	−	−	+	+	+	−	−	
−	+	−	+	+	−	−	+	Block 2
+	+	+	+	−	+	+	−	
				−	+	−	+	Block 3
				+	−	+	−	
				+	+	+	+	Block 4
				−	−	−	−	

to that obtained earlier, and illustrated in Table 12, in the discussion of blocking designs of resolution III. Similarly, the 2_{IV}^{8-4} design in blocks of two obtained in Table 20b could also have been formed by using the principle of fold-over starting out with the 2_{III}^{7-3} along with a column vector of plus signs and pairing each run with its fold-over. The same is true of the 2_{IV}^{16-11} obtained by fold-over described in Table 17b. The designs obtained by folding over the Plackett and Burman designs can also be broken into blocks of two runs each using precisely the same device.

In general, any resolution IV design provides an opportunity to obtain blocks of size two whose effects do not confound any of the main effects. In doing this, of course, we confound the two-factor interactions with blocks. Nevertheless, the resulting designs are of considerable interest. Often, the comparisons between blocks merely represent influences upon the response having a somewhat higher variation than that responsible for differences within blocks. In these circumstances it is reasonable to think of the strings of two-factor interactions simply as being estimated with a variance somewhat higher than that appropriate for the main effects. For instance, these designs may be used where it is suspected that a time trend may occur during the course of the trials. Provided proper randomization is applied both to the order of runs within the blocks and to the order of running the blocks themselves, the design is such that whereas main effects were determined with a variance appropriate to successive observations, the strings of two-factor interactions would be estimated with a variance appropriate to pairs of runs made in random order in the presence of a time trend.

"Major and "Minor" Variables in Resolution IV *Designs*

We have already seen that a 2^{p-q} design of resolution IV can be regarded from two points of view: (1) it is a design suitable for providing estimates of the p main effects even though two-factor interactions may occur, and (2) it is a design suitable for providing unbiased estimates of all main effects and interactions between *any* three of the factors if the others are of no importance. The designs can be considered from still another point of view. Considering the 2_{IV}^{16-11} design as an example, we have seen how this design may be run in sixteen blocks of two runs each, the blocks being obtained from four block generators associated with two-factor interactions. Alternatively we can choose the four block generators to represent actual variables. Suppose for example, we have four "major" variables for which we wish to estimate all the main effects and all the interactions and we have sixteen further variables which we believe exert at most main effects, and may be conveniently viewed as "minor" variables. Then this design may be employed associating the four major variables with the block generators and the remaining minor variables with the sixteen "main effect" factors. Of course, all the effects among the major variables will now be confounded with the sets of two-factor interactions of the minor variables. However, since minor variables are believed to exert at most main effects, the two-factor interactions between these variables are tentatively assumed to be nil.

In this connection, there is an opportunity to make use of any prior feeling which the experimenter may have concerning the possibility of interaction in the minor variables. Should he feel particularly anxious about a possible interaction between two minor variables, then he can usually arrange, by inspecting tables such as that given in Equation (26), that this interaction is associated with an unimportant interaction between the major variables. For instance, in this present example, the interactions **1 6**, **7 8**, **3 4** and **2 5** between the minor variables are all confounded with the three-factor interaction $\mathbf{B}_1\mathbf{B}_2\mathbf{B}_3$ between major variables. This three factor interaction might be expected to be unimportant *a priori*. It should be noticed that so long as $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_3$ are pseudo variables representing comparisons among blocks, then interactions such as $\mathbf{B}_1\mathbf{B}_2\mathbf{B}_3$ will represent comparisons of precisely the same potential as are represented by the main effects $\mathbf{B}_1$, $\mathbf{B}_2$, etc. When, however, $\mathbf{B}_1$, $\mathbf{B}_2$, etc., are used to represent real variables, main effects and interactions revert to their former relative status.

When thirty-two runs are to be made and where there are four major variables along with sixteen minor variables, the 2_{IV}^{16-11} design may be employed. With sixteen runs three variables and all their interactions may be investigated by associating the block generators with these major variables and the eight minor variables then introduced. For eight runs, two major variables and their interaction plus four main effect variables are possible. Of course, when the designs are used in this way no blocking is permissible. However, even here a certain degree of flexibility is possible. For instance, for the thirty-two run design we might wish to have only two principal factors in which case we could associate these with two of the block generators using the other two block generators to form blocks of eight. It will be clear to the reader that these arrangements provide a very versatile set of designs which may be used in a variety of circumstances.

From Resolution IV *Designs to Resolution* III *Designs*

When the 2_{IV}^{16-11} designs is used to study simultaneously four major variables along with sixteen minor variables, a convenient notation for the design is

$$2^4 \subset 2_{IV}^{16-11}$$

where the symbol $\subset$ is read "contained in" or "embedded in". Thus $2^3 \subset 2_{IV}^{8-4}$ and $2^2 \subset 2_{IV}^{4-1}$ identify the sixteen and eight run designs described above.

The construction of a resolution III design from one of resolution IV now becomes obvious. The $2^4 \subset 2_{IV}^{16-11}$ is clearly a design for studying twenty variables in thirty-two runs. Suppose now that one of the interactions between the major variables is used to bring in still another variable. In fact, we might be willing to assume that all the interactions between the four major variables are negligible and in this instance eleven new variables (one for each of the interactions) could be introduced. The result, of course, is the 2_{III}^{31-26} design, that is, the fully saturated resolution III design for studying thirty-one variables in thirty-two runs.

Other Embedded Designs

The principle of embedded fold-over pairs producing blocks of two runs has wide application. As an example less orthodox than those mentioned above, we note that the thirty-two run 2_{IV}^{16-11} design in sixteen blocks is one in which a central composite design in three variables can be embedded. The composite design [14] would employ factor combinations of three major variables consisting of a 2^3 factorial along with six axial points and two center points for a total of sixteen points. Each of these factor combinations would then be duplicated, one duplicate containing one combination of versions of sixteen minor variables with the second duplicate the mirror image of these same sixteen variables. The additional sixteen variables would have to be such that they were not expected to have any effect other than a linear one.

Part II of this paper will contain a discussion of Resolution V designs along with an appendix.

Bibliography

(1) Yates, F., The Design and Analysis of Factorial Experiments., *Imper. Bur. Soil Sci. Tech. Comm.* 35, 1937.

(2) Fisher, R. A., The Theory of Confounding in Factorial Experiments, in Relation to the Theory of Groups, *Ann. Eugen.* 11, 1942.

(3) Daniel, C., Fractional Replication in Industrial Research, *Proc. 3rd Berkeley Symp.* V., 1956.

(4) Finney, D. J., The Fractional Replication of Factorial Arrangements, *Ann. Eugen.* 12, 1945.

(5) Brownlee, K. A., B. K. Kelly and P. K. Loraine, Fractional Replication Arrangements for Factorial Experiments with Factors at Two Levels, *Biometrika*, 35, 1948.

(6) Plackett, R. L., and J. P. Burman, The Design of Optimum Multifactorial Experiments, *Biometrika*, 33, 1946.
(7) Clatworthy, W. H., W. S. Connor, Deming and M. Zelen, Fractional Factorial Experiment Designs for Factors at Two Levels, U. S. *Dept. Comm. Nat. Bur. of Stnds., Applied Math.* Ser. 48, 1957.
(8) Kempthorne, O., A Simple Approach to Confounding and Fractional Replication in Factorial Experiments, *Biometrika*, 34, 1947.
(9) Bose, R. C. and Kishen, K., On the Problem of Confounding in the General Symmetrical Factorial Design, *Sankhya* 5, 1940.
(10) Dykstra, O., Partial Duplication of Factorial Experiments, *Technometrics*, 1, 1959.
(11) Plackett, R. L., Some Generalizations in the Multifactorial Design, *Biometrika*, 33, 1946.
(12) Fisher, R. A., *The Design of Experiments*, Oliver & Boyd, London, 1935.
(13) Box, G. E. P., Multifactor Designs of First Order, *Biometrika*, 39, 1951.
(14) Box, G. E. P. & K. B. Wilson, On the Experimental Attainment of Optimum Condition, *J. Roy. Stat. Soc., B.*, 13, 1951.
(15) Tippett, L. H. C., "Applications of Statistical Methods to the Control of Quality in Industrial Production", Manchester Statistical Society, 1934.
(16) Yates, F., Complex Experiments, *J. Roy. Stat. Soc., Suppl.*, 2, 1935.
(17) Rao, C. R., Factorial Experiments Derivable from Combinatorial Arrangements of Arrays, *J. Roy. Stat. Soc., Suppl.*, 9, 1947.