Designation: E1325 – 16

An American National Standard

Standard Terminology Relating to Design of Experiments[1]

This standard is issued under the fixed designation E1325; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This standard includes those statistical items related to the area of design of experiments for which standard definitions appear desirable.

2. Referenced Documents

2.1 *ASTM Standards:*[2]

E456 Terminology Relating to Quality and Statistics

3. Significance and Use

3.1 This standard is a subsidiary to Terminology E456.

3.2 It provides definitions, descriptions, discussion, and comparison of terms.

4. Terminology

aliases, *n—in a fractional factorial design*, two or more effects which are estimated by the same contrast and which, therefore, cannot be estimated separately.

DISCUSSION—(*1*) The determination of which effects in a 2^n factorial are *aliased* can be made once the *defining contrast* (in the case of a half replicate) or *defining contrasts* (for a fraction smaller than ½) are stated. The *defining contrast* is that effect (or effects), usually thought to be of no consequence, about which all information may be sacrificed for the experiment. An identity, *I*, is equated to the *defining contrast* (or *defining contrasts*) and, using the conversion that $A^2 = B^2 = C^2 = I$, the multiplication of the letters on both sides of the equation shows the aliases. In the example under fractional factorial design, I = ABCD. So that: $A = A^2BCD = BCD$, and $AB = A^2B^2CD = CD$.

(*2*) With a large number of factors (and factorial treatment combinations) the size of the experiment can be reduced to ¼, ⅛, or in general to $½^{k}$ to form a 2^{n-k} fractional factorial.

(*3*) There exist generalizations of the above to factorials having more than 2 levels.

balanced incomplete block design (BIB), *n*—an incomplete block design in which each block contains the same number k of different versions from the t versions of a single principal factor arranged so that every pair of versions occurs together in the same number, λ, of blocks from the b blocks.

DISCUSSION—The design implies that every version of the principal factor appears the same number of times r in the experiment and that the following relations hold true: $bk = tr$ and $r(k - 1) = \lambda(t - 1)$.

For randomization, arrange the blocks and versions within each block independently at random. Since each letter in the above equations represents an integer, it is clear that only a restricted set of combinations (t, k, b, r, λ) is possible for constructing balanced incomplete block designs. For example, $t = 7$, $k = 4$, $b = 7$, $\lambda = 2$. Versions of the principal factor:

Block					
	1	1	2	3	6
	2	2	3	4	7
	3	3	4	5	1
	4	4	5	6	2
	5	5	6	7	3
	6	6	7	1	4
	7	7	1	2	5

block factor, *n*—a factor that indexes division of experimental units into disjoint subsets.

DISCUSSION—Blocks are sets of similar experimental units intended to make variability within blocks as small as possible, so that treatment effects will be more precisely estimated. The effect of a block factor is usually not of primary interest in the experiment. Components of variance attributable to blocks may be of interest. The origin of the term "block" is in agricultural experiments, where a block is a contiguous portion of a field divided into experimental units, "plots," that are each subjected to a treatment.

completely randomized design, *n*—a design in which the treatments are assigned at random to the full set of experimental units.

DISCUSSION—No block factors are involved in a completely randomized design.

completely randomized factorial design, *n*—a factorial experiment (including all replications) run in a completely randomized design.

composite design, *n*—a design developed specifically for fitting second order response surfaces to study curvature, constructed by adding further selected treatments to those obtained from a 2^n factorial (or its fraction).

DISCUSSION—If the coded levels of each factor are – 1 and + 1 in the 2^n factorial (see notation 2 under discussion for **factorial experiment**), the $(2n + 1)$ additional combinations for a *central composite design* are

[1] This terminology is under the jurisdiction of ASTM Committee E11 on Quality and Statistics and is the direct responsibility of Subcommittee E11.10 on Sampling / Statistics.

Current edition approved April 1, 2016. Published April 2016. Originally approved in 1990. Last previous edition approved in 2015 as E1325 – 15. DOI: 10.1520/E1325-16.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

(0, 0, ..., 0), ($\pm a$, 0, 0, ..., 0) 0, $\pm a$, 0, ..., 0) ..., (0, 0, ..., $\pm$ a). The minimum total number of treatments to be tested is ($2^n + 2n + 1$) for a 2^n factorial. Frequently more than one center point will be run. For $n = 2$, 3 and 4 the experiment requires, 9, 15, and 25 units respectively, although additional replicate runs of the center point are usual, as compared with 9, 27, and 81 in the 3^n factorial. The reduction in experiment size results in confounding, and thereby sacrificing, all information about curvature interactions. The value of a can be chosen to make the coefficients in the quadratic polynomials as orthogonal as possible to one another or to minimize the bias that is created if the true form of response surface is not quadratic.

confounded factorial design, *n*—a factorial experiment in which only a fraction of the treatment combinations are run in each block and where the selection of the treatment combinations assigned to each block is arranged so that one or more prescribed effects is(are) confounded with the block effect(s), while the other effects remain free from confounding.

NOTE 1—All factor level combinations are included in the experiment.

DISCUSSION—*Example:* In a 2^3 factorial with only room for 4 treatments per block, the *ABC* interaction (*ABC*: $-(1) + a + b - ab + c - ac - bc + abc$) can be sacrificed through confounding with blocks without loss of any other effect if the blocks include the following:

	Block 1	Block 2
Treatment	(1)	*a*
Combination	*ab*	*b*
(Code identification shown in discussion under factorial experiment)	*ac*	*c*
	bc	*abc*

The treatments to be assigned to each block can be determined once the effect(s) to be confounded is(are) defined. Where only one term is to be confounded with blocks, as in this example, those with a positive sign are assigned to one block and those with a negative sign to the other. There are generalized rules for more complex situations. A check on all of the other effects (*A, B, AB*, etc.) will show the balance of the plus and minus signs in each block, thus eliminating any confounding with blocks for them.

confounding, *n*—combining indistinguishably the main effect of a factor or a differential effect between factors (interactions) with the effect of other factor(s), block factor(s) or interaction(s).

NOTE 2—Confounding is a useful technique that permits the effective use of specified blocks in some experiment designs. This is accomplished by deliberately preselecting certain effects or differential effects as being of little interest, and arranging the design so that they are confounded with block effects or other preselected principal factor or differential effects, while keeping the other more important effects free from such complications. Sometimes, however, confounding results from inadvertent changes to a design during the running of an experiment or from incomplete planning of the design, and it serves to diminish, or even to invalidate, the effectiveness of an experiment.

contrast, *n*—a linear function of the observations for which the sum of the coefficients is zero.

NOTE 3—With observations Y_1, Y_2, ..., Y_n, the linear function $a_1Y_1 + a_2Y_2 + ... + a_1Y_n$ is a contrast if, and only if $\sum a_i = 0$, where the a_i values are called the contrast coefficients.

DISCUSSION—*Example 1:* A factor is applied at three levels and the results are represented by A_1, A_2, A_3. If the levels are equally spaced, the first question it might be logical to ask is whether there is an overall linear trend. This could be done by comparing A_1 and A_3, the extremes of A in the experiment. A second question might be whether there is evidence that the response pattern shows curvature rather than a simple linear trend. Here the average of A_1 and A_3 could be compared to A_2. (If there is no curvature, A_2 should fall on the line connecting A_1 and A_3 or, in other words, be equal to the average.) The following example illustrates a regression type study of equally spaced continuous variables. It is frequently more convenient to use integers rather than fractions for contrast coefficients. In such a case, the coefficients for Contrast 2 would appear as (−1, + 2, − 1).

Response	A_1	A_2	A_3
Contrast coefficients for question 1	−1	0	+1
Contrast 1	$-A_1$	...	$+ A_3$
Contrast coefficients for question 2	−½	+1	−½
Contrast 2	−½ A_1	$+ A_2$	−½ A_3

Example 2: Another example dealing with discrete versions of a factor might lead to a different pair of questions. Suppose there are three sources of supply, one of which, A_1, uses a new manufacturing technique while the other two, A_2 and A_3 use the customary one. First, does vendor A_1 with the new technique seem to differ from A_2 and A_3? Second, do the two suppliers using the customary technique differ? Contrast A_2 and A_3. The pattern of contrast coefficients is similar to that for the previous problem, though the interpretation of the results will differ.

Response	A_1	A_2	A_3
Contrast coefficients for question 1	−2	+1	+1
Contrast 1	$-2A_1$	$+A_2$	$+A_3$
Contrast coefficients for question 2	0	−1	+1
Contrast 2	...	$- A_2$	$+ A_3$

The coefficients for a contrast may be selected arbitrarily provided the $\sum a_i = 0$ condition is met. Questions of logical interest from an experiment may be expressed as contrasts with carefully selected coefficients. See the examples given in this discussion. As indicated in the examples, the response to each treatment combination will have a set of coefficients associated with it. The number of linearly independent contrasts in an experiment is equal to one less than the number of treatments. Sometimes the term *contrast* is used only to refer to the pattern of the coefficients, but the usual meaning of this term is the algebraic sum of the responses multiplied by the appropriate coefficients.

contrast analysis, *n*—a technique for estimating the parameters of a model and making hypothesis tests on preselected linear combinations of the treatments (contrasts). See Table 1 and Table 2.

NOTE 4—Contrast analysis involves a systematic tabulation and analysis format usable for both simple and complex designs. When any set of orthogonal contrasts is used, the procedure, as in the example, is straightforward. When terms are not orthogonal, the orthogonalization process to adjust for the common element in nonorthogonal contrast is also systematic and can be programmed.

DISCUSSION—*Example:* Half-replicate of a 2^4 factorial experiment with factors *A, B* and *C* (X_1, X_2 and X_3 being quantitative, and factor *D* (X_4) qualitative. Defining contrast $I = + ABCD = X_1X_2X_3X_4$ (see **fractional factorial design** and **orthogonal contrasts** for derivation of the contrast *coefficients*).

design of experiments, *n*—the arrangement in which an experimental program is to be conducted, and the selection of the levels (versions) of one or more factors or factor combinations to be included in the experiment. Synonyms include experiment design and experimental design.

DISCUSSION—The purpose of designing an experiment is to provide the most efficient and economical methods of reaching valid and relevant conclusions from the experiment. The selection of an appropriate design for any experiment is a function of many considerations such as the type of questions to be answered, the degree of generality to be attached to the conclusions, the magnitude of the effect for which a high probability of detection (power) is desired, the homogeneity of the experimental units and the cost of performing the experiment. A

TABLE 1 Contrast Coefficient

Source	Treatments	(1)	*ab*	*ac*	*bc*	*ad*	*bd*	*cd*	*abcd*	
Centre	X_0	+1	+1	+1	+1	+1	+1	+1	+1	See Note 1
A(+*BCD*): pH (8.0; 9.0)	X_1	−1	+1	+1	−1	+1	−1	−1	+1	
B(+ *ACD*): SO_4 (10 cm^3; 16 cm^3)	X_2	−1	+1	−1	+1	−1	+1	−1	+1	
C(+ *ABD*): Temperature (120°C; 150°C)	X_3	−1	−1	+1	+1	−1	−1	+1	+1	
D(+*ABC*): Factory (P; Q)	X_4	−1	−1	−1	−1	+1	+1	+1	+1	
AB + *CD*	$X_1X_2 = X_{12}$	+1	+1	−1	−1	−1	−1	+1	+1	
AC + *BD*	$X_1X_3 = X_{13}$	+1	−1	+1	−1	−1	+1	−1	+1	See Note 2
AD + *BC*	$X_1X_4 = X_{14}$	+1	−1	−1	+1	+1	−1	−1	+1	

NOTE 1—The center is not a constant ($\sum X_i \neq 0$) but is convenient in the contrast analysis calculations to treat it as one.

NOTE 2—Once the contrast coefficients of the main effects (X_1, X_2, X_3 and X_4) are filled in, the coefficients for all interaction and other second or higher order effects can be derived as products ($X_{ij} = X_i X_i$) of the appropriate terms.

TABLE 2 Contrast Analysis

Source	Contrast $\sum_i X_{ij} Y_i^1$	Divisor $\sum_i X_{ij}^2$	Student's *t* ratio[2] $(\sum_i X_{ij} Y_i)/s\sqrt{\sum_i X_{ij}^2}$	Regression coefficient $B_j = (\sum_i X_{ij} Y_I) / \sum_i X_{ij}^2$
X_0: Centre	$\sum X_0 Y$	$\sum X_0^2$	$(\sum X_0 Y)/s\sqrt{\sum X_0^2}$	$B_0 = (\sum X_0 Y)/\sum X_0^2$
X_1: *A* + *BCD*	$\sum X_1 Y$	$\sum X_1^2$	$(\sum X_1 Y)/s\sqrt{\sum X_1^2}$	$B_1 = (\sum X_1 Y)/\sum X_1^2$
X_2: *B* + *ACD*	$\sum X_2 Y$	$\sum X_2^2$	$(\sum X_2 Y)/s\sqrt{\sum X_2^2}$	$B_2 = (\sum X_2 Y)/\sum X_2^2$
X_3: *C* + *ABD*	$\sum X_3 Y$	$\sum X_3^2$	$(\sum X_3 Y)/s\sqrt{\sum X_3^2}$	$B_3 = (\sum X_3 Y)/\sum X_3^2$
X_4: *D* + *ABC*	$\sum X_4 Y$	$\sum X_4^2$	$(\sum X_4 Y)/s\sqrt{\sum X_4^2}$	$B_4 = (\sum X_4 Y)/\sum X_4^2$
X_{12}: *AB* + *CD*	$\sum X_{12} Y$	$\sum X_{12}^2$	$(\sum X_{12} Y)/s\sqrt{\sum X_{12}^2}$	$B_{12} = (\sum X_{12} Y)/\sum X_{12}^2$
X_{13}: *AC* + *BD*	$\sum X_{13} Y$	$\sum X_{13}^2$	$(\sum X_{13} Y)/s\sqrt{\sum X_{13}^2}$	$B_{13} = (\sum X_{13} Y)/\sum X_{13}^2$
X_{14}: *AD* + *BC*	$\sum X_{14} Y$	$\sum X_{14}^2$	$(\sum X_{14} Y)/s\sqrt{\sum X_{14}^2}$	$B_{14} = (\sum X_{14} Y)/\sum X_{14}^2$

NOTE 1—The notation for contrast analysis usually uses *Y* to indicate the response variable and *X* the predictor variables.

NOTE 2—The measure of experimental error, *s*, can be obtained in various ways. If the experiment is replicated, *s* is the square root of the pooled variances of the pairs for each treatment combination. (Each row of *X* values would be expanded to account for the additional observations in the contrast analysis computations). If some effects were felt to be pseudo-replicates (example, no interactions were logical) multiplying the contrast by the regression coefficient of these terms forms a sum of squares (as in analysis of variance) and these would be summed and divided by the number of terms involved to give s^2. Also, in many experiments, past experience may already provide an estimate of this error. Assumed model: $Y = B_0 + B_1X_{1i} + B_2X_{3i} + B_4X_{4i} + e$). In a simple 2-level experiment such as this, the regression coefficient measures the half-effect of shifting a factor, say pH, between its low and high level, or the effect of shifting from a center level to the high level. In general, substitution of the appropriate contrast coefficients for the *X* terms in the model will permit any desired comparisons. The difference between quantitative and qualitative factors lies in the interpretation. Since a unit of X_1 represents a pH shift of 0.5, there is a meaningful translation into physical units. On the other hand, the units of the qualitative variable (factories) have no significance other than for identification and in the substitution process to obtain estimates of the average response values.

properly designed experiment will permit relatively simple statistical interpretation of the results, which may not be possible otherwise. The *arrangement* includes the randomization procedure for allocating treatments to experimental units.

experimental design, *n*—see **design of experiments.**

experimental unit, *n*—a portion of the experiment space to which a treatment is applied or assigned in the experiment.

NOTE 5—The unit may be a patient in a hospital, a group of animals, a production batch, a section of a compartmented tray, etc.

experiment space, *n*—the materials, equipment, environmental conditions and so forth that are available for conducting an experiment.

DISCUSSION—That portion of the experiment space restricted to the range of levels (versions) of the factors to be studied in the experiment is sometimes called the *factor space*. Some elements of the experiment space may be identified with blocks and be considered as block factors.

evolutionary operation (EVOP), *n*— a sequential form of experimentation conducted in production facilities during regular production.

NOTE 6—The principal theses of EVOP are that knowledge to improve the process should be obtained along with a product, and that designed experiments using relatively small shifts in factor levels (within production tolerances) can yield this knowledge at minimum cost. The range of variation of the factors for any one EVOP experiment is usually quite small in order to avoid making out-of-tolerance products, which may require considerable replication, in order to be able to clearly detect the effect of small changes.

factor, *n*—independent variable in an experimental design.

DISCUSSION—Factors can include controllable factors that are of interest for the experiment, block factors that are created to enhance precision of the factors of interest, and uncontrolled factors that might be measured in the experiment. Design of an experiment consists of allocating levels of each controllable experimental factor to experimental units.

2^n factorial experiment, *n*—a factorial experiment in which *n* factors are studied, each of them in two levels (versions).

DISCUSSION—The 2^n factorial is a special case of the general factorial. (See **factorial experiment (general).**) A popular code is to indicate a small letter when a factor is at its high level, and omit the letter when it is at its low level. When factors are at their low level the code is (*1*).

Example (illustrating the discussion)—A2^3 factorial with factors *A, B*, and *C*:

	Level							
Factor A	Low	High	Low	High	Low	High	Low	High
Factor B	Low	Low	High	High	Low	Low	High	High
Factor C	Low	Low	Low	Low	High	High	High	High
Code	(*1*)	*a*	*b*	*ab*	*c*	*ac*	*bc*	*abc*

This type of identification has advantages for defining blocks, confounding and aliasing. See **confounded factorial design** and **fractional factorial design**.

Factorial experiments regardless of the form of analysis used, essentially involve contrasting the various levels (versions) of the factors.

Example (illustrating contrast)—Two-factor, two-level factorial 2^2 with factors *A* and *B*: *A* = [a – (1)] + [*ab* – *b*]. This is the contrast of *A* at the *low* level of *B* plus the contrast of *A* at the *high* level of *B*. *B* = [*b* – (1)] + [*ab* – *a*]. This is the contrast of *B* at the *low* level of *A* plus the contrast of *B* at the *high* level of *A: AB* = [*ab* – *b*] – [*a* – (1)] = [*ab* – *a*] – [*b* – (*1*)]. This is the contrast of the contrasts of *A* at the *high* level of *B* and the *low* level of *B* or the contrast of the contrasts of *B* at the *high* level of *A* and at the *low* level of *A*.

Each contrast can be derived from the development of a *symbolic product* of two factors, these factors being of the form ($a \pm 1$), ($b \pm 1$), using – 1 when the capital letter (*A, B*) is included in the contrast and + 1 when it is not.

Example:

$$A: (a-1)(b+1)$$
$$B: (a+1)(b-1)$$
$$AB: (a-1)(b-1)$$

These expressions are usually written in a standard order, in this case:

$$A: -(1) + a - b + ab$$
$$B: -(1) - a + b + ab$$
$$AB: (1) - a - b + ab$$

Note that the coefficient of each treatment combination in AB (+ 1 or – 1) is the product of the corresponding coefficients in A and B. This property is general in 2^n factorial experiments. After grouping, the A term 2^n represents the effect of A averaged over the two levels of B, that is, a main effect or average effect. Similarly, B represents the average effect of B over both levels of A. The AB term contrasts the effect of A at the high and the low levels of B (or the effect of B at the high and low levels of A), that is an interaction or differential effect.

This example is, of course, the simplest case, but it illustrates the basic principles. The contrasts may appear more complex as additional factors are introduced.

factorial experiment (general), *n*—in general, an experiment in which all possible treatments formed from two or more factors, each being studied at two or more levels (versions), are examined so that interactions (differential effects) as well as main effects can be estimated.

DISCUSSION—The term is descriptive of the combining of the various factors in all possible combinations, but in itself does not describe the experimental design in which these combinations, or a subset of these combinations, will be studied.

The most commonly used designs for the selected arrangement of the factorial treatment combinations are the completely randomized design, the randomized block design and the balanced incomplete block design, but others also are used.

A factorial experiment is usually described symbolically as the product of the number of levels (versions) of each factor. For example, an experiment based on 3 levels of factor *A*, 2 versions of factor *B* and 4 levels of factor *C* would be referred to as a $3 \times 2 \times 4$ factorial. The product of these numbers indicates the number of factorial treatments.

When a factorial experiment includes factors all having the same number of levels (versions), the description is usually given in terms of the number of levels raised to the power equal to the number of factors, *n*. Thus, an experiment with three factors all run at two levels would be referred to as a 2^3 factorial (*n* being equal to 3) and has 8 factorial treatment combinations. Some commonly used notations for describing the treatment combinations for a factorial experiment are as follows:

(1) Use a letter to indicate the factor and a numerical subscript the level (version) of the factor, for example, three factors *A, B*, and *C* in a $2 \times 3 \times 2$ factorial. The 12 combinations would be:

$$A_1B_1C_1,\ A_2B_1C_1,\ A_1B_2C_1,\ A_2B_2C_1,\ A_1B_3C_1,\ A_2B_3C_1,$$
$$A_1B_1C_2,\ A_2B_1C_2,\ A_1B_2C_2,\ A_2B_2C_2,\ A_1B_3C_2,\ A_2B_3C_2.$$

Sometimes only the subscripts, listed in the same order as the factors are used, such as: 111, 211, 121, 221, 131, 231, 112, 212, 122, 222, 132, 232. A variation which permits the use of modulo 2 and modulo 3 arithmetic for the purpose of listing the treatment combinations in blocked and fractional designs is: 000, 100, 010, 110, 020, 120, 001, 101, 011, 111, 021, 121.

(*2*) Describe the levels in terms of the number of unit deviations from the center level, including sign. In the case of an even number of levels where there is no actual treatment at the center level, the coefficients describing the levels are usually given in terms of half-unit deviations. For example, with two levels, if a unit of deviation between these levels is 4 mm, the –1 coefficient might be assigned to 3 mm and the + 1 to 7 mm with 0 being assigned to the non-included 5 mm level. In this example the code would appear as follows.

(–1, –1, –1); (+1, –1, –1); (–1, 0, –1); (+1, 0, –1);
(–1, +1, –1); (+1, +1, –1); (–1, –1, +1); (+1, –1, +1);
(–1, 0, +1); (+1, 0, +1); (–1, +1, +1); (+1, +1, +1)

This descriptive coding has many advantages, particularly in analyzing contrasts when levels are equally spaced. Unequal spacing of the levels or weighted emphasis for the various versions can also be reflected in the coefficients.

fractional factorial design, *n*—a factorial experiment in which only an adequately chosen fraction of the treatments required for the complete factorial experiment is selected to be run.

NOTE 7—This procedure is sometimes called fractional replication.

DISCUSSION—The fraction selected is obtained by choosing one or several *defining contrasts* which are considered of minor importance, or negligible, generally interaction(s) of high order. These *defining contrasts* cannot be estimated and thus are sacrificed. By *adequately chosen* is meant selection according to specified rules which include consideration of effects to be confounded and aliased (see **confounding** and **aliases**). It is possible to use tables of orthogonal arrays, algorithms or a listing of designs to obtain the factorial treatment combinations for the fractional replicate without actually specifying the defining contrasts, but this entails a loss of information.

Fractional factorial designs are often used very effectively in screening tests to determine which factor or factors are large contributors to variability, or as part of a sequential series of tests, but there are risks of getting biased estimates of main effects or of misjudging the relative importance of various factors. When there is a large number of factor level combinations resulting from a large number of factors to be tested, it is often impracticable to test all the combinations with one experiment. In such cases resort may be made to a fractional, that is, partial, replication. The usefulness of these designs stems from the fact that, in general, higher order interactions are not likely to occur. When this assumption is not valid, biased estimates will result.

Example—Two half-replicates of a 2^4 factorial (refer to the discussion under factorial experiment for the code interpretation). Defining contrast, *ABCD*:

+	–
abcd	*abc*
ab	*abd*
ac	*acd*
ad	*bcd*
bc	*a*
bd	*b*
cd	*c*
(1)	*d*

Either of these half-replicates can be used as a *fractional replicate.*

In the example, the factorial combinations in the first column are those with *a* + (plus) sign in the development of symbolic product of the *ABCD defining contrast*, as illustrated in the example o f 2^n factorial experiment. *ABCD* = (*a* – 1)(*b* – 1)(*c* – 1)(*d* – 1). Those factorial combinations in the second column are those with *a* – (minus) sign.

Because only those elements of the *ABCD* interaction having the same sign are run, no *ABCD* contrast measure is obtainable, so that the *ABCD* interaction is completely confounded and unestimable. In addition, it will be found that because only half of the full factorial experiment is run, each contrast represents two effects.

From the + sign fractional replicate in this example, we should compute the factorial effects as follows:

$$A = (abcd)+(ab)+(ac)+(ad)-(bc)-(bd)-(cd)-(1) = BCD$$
$$AB = (abcd)+(ab)+(cd)+(1)-(ac)-(ad)-(bc)-(bd) = CD$$

Effects represented by the same contrast are named *aliases*. (See **aliases**). Note that had the complete set of factorial treatments been run instead of only half of them, the estimates of the *A* and *BCD* or *AB* and *BC* effects would no longer be identical. That is, when all 16 combinations are included instead of only 8:

$$A = (a-1)(b+1)(c+1)(d+1) \quad (1)$$

is not equal to:

$$BCD = (a+1)(b-1)(c-1)(d-1) \quad (2)$$

This example, and the comments thereon, have been limited to the 2^n factorial experiments. A comparable, but more difficult, approach is available when there are more than two versions, but another approach to these situations is through the use of the composite design.

fully nested experiment—a nested experiment in which the second factor is nested within levels (versions) of the first factor and each succeeding factor is nested within versions of the previous factor.

DISCUSSION—*Example:*

Factor *A* version	A_1				A_2			
Factor *B* version	B_1		B_2		B_3		B_4	
Factor *C* version	C_1	C_2	D_3	C_4	C_5	C_6	C_7	C_8

hierarchical experiment, *n*—see **nested experiment**.

incomplete block design, *n*—a design in which the experiment space is subdivided into blocks in which there are insufficient experimental units available to run a complete set of treatments or replicate of the experiment.

interaction, *n*—differences in responses to a factor among levels (versions) of other factors in the experiment.

DISCUSSION—When factors do not interact, the joint effect on the experiment response can be modelled as a sum of level effects for each factor. Interacting factors are such that the effect on the response with respect to one variable also depends on the level of one or more other variables.

latin square, *n*—a factorial experiment having two block factors (rows and columns) and a treatment factor, with equal numbers of levels, and for which each treatment occurs once in each row and column.

DISCUSSION—*Example:* Five treatments A, B, C, D, E in a 5 × 5 latin square.

		Column				
		1	2	3	4	5
Row	1	D	A	E	B	C
	2	E	B	A	C	D
	3	A	C	B	D	E
	4	B	D	C	E	A
	5	C	E	D	A	B

level (of a factor), *n*—a given value, a specification of procedure or a specific setting of a factor.

NOTE 8—*Version* is a general term applied both to quantitative and qualitative factors. The more restrictive term *level* is frequently used to express more precisely the quantitative characteristic. For example, two versions of a catalyst may be presence and absence. Four levels of a heat treatment may be 110°C, 120°C, 140°C, and 160°C.

DISCUSSION—Responses observed at the various levels (versions) of a factor provide information for determining the effect of the factor within the range of levels of the experiment. Extrapolation beyond the range of these levels is usually inappropriate without a firm basis for assuming model relationships. Interpolation within the range may depend on the number of levels and the spacing of these levels. It is usually reasonable to interpolate, although it is possible to have discontinuous or multimodal relationships that cause abrupt changes within the range of the experiment. The levels (versions) may be limited to certain selected fixed values (whether these values are or are not known) or they may represent purely random selection over the range to be studied. The method of analysis is dependent on this selection.

main effect, average effect, *n*—a term describing a measure for the comparison of the responses at each level (version) of a factor averaged over all levels (versions) of other factors in the experiment.

NOTE 9—The term *main effect* may describe the parameter in an assumed model or the estimate of this parameter.

DISCUSSION—It should be noted that even though a main effect is indicated to be small, this does not necessarily mean that the factor is unimportant. Large effects of the factor may result at various levels (versions) of other factors, but may differ in sign or magnitude or both. The process of averaging in these cases would tend to make the main effect appear smaller. See **interaction**.

method of least squares, *n*—a technique of estimation of a parameter which minimizes $\sum e^2$, where *e* is the difference between the observed value and the predicted value derived from the assumed model.

DISCUSSION—The experimental errors associated with the individual observations ordinarily are assumed to be independent, although the method may be generalized to the case of correlated errors. The usual analysis of variance, regression analysis and contrast analysis are all based on the method of least squares and provide different computational and interpretative advantages that stem from certain balances within the experimental arrangements which permit convenient groupings of the data.

mixture design, *n*—a design in which two or more ingredients or components shall be mixed and the response is a property of the resulting mixture that does not depend upon the amount of the mixture.

NOTE 10—The proportions of each of the *q* components (X_i) in the

mixture shall satisfy the conditions $O \le X_i \le 1$ and $\sum_{i=1}^{c} X_i = 1$; and each experimental point is defined in terms of these proportions.

NOTE 11—In some fields of application the experiment mixtures are described by the terms *formulation* or *blend*. The use of mixture designs is appropriate for experimenting with the formulations of manufactured products, such as paints, gasoline, foods, rubber, and textiles.

NOTE 12—In some applications, the proportions of the components of the mixture may vary between 0 and 100 % of the mixture (*complete domain*). In others, there may be operative restraints, so that at least one component cannot attain 0 or 100 % (*reduced domain*).

nested experiment, *n*—an experiment to examine the effect of two or more factors in which the same level (version) of a factor cannot be used with all levels (versions) of other factors. Synonym: hierarchical experiment.

DISCUSSION—Generally, nested experiments are used to evaluate studies in terms of components of variance rather than in terms of differences in response levels or prediction models. See the discussion under Model 2 analysis of variance.

Example:

Vendor	A		B	
Shipment	1	2	3	4

If two vendors are to be compared by evaluating two shipments from each, there ordinarily is no direct relationship between the first shipment of Vendor A and that of Vendor B or similarly for the second shipment. The differences between the two versions of the shipment factor of Vendor A are nested within that version of the vendor factor and, similarly, the differences between the two versions of the shipment factor of Vendor B are nested within this other version of the vendor factor.

It is sometimes possible to redefine the factor into versions that can be compared across other factors if that makes a more meaningful question. For example, Shipments 1 and 3 of the above example might represent Monday morning production and Shipments 2 and 4 Friday afternoon production. The question could be framed in terms of Monday morning versus Friday afternoon production, which has a common thread, rather than in terms of two unrelated shipments. This would now represent a crossed [for example each level (version) of a factor is used with all levels (versions) of the other factors], rather than nested, classification and could be arranged as a factorial experiment.

	Vendor	A	B
Day	Monday	1	1
	Friday	2	2

orthogonal array, *n*—a table of coefficients identifying the levels, or some weight associated with the levels, for each factor to be used in the analysis of specified effects, which are arranged in such a manner that each effect will be independent of the other effects.

DISCUSSION—Listings of some orthogonal arrays may be found in most textbooks on the design of experiments under the headings of fractional factorial designs, latin squares, orthogonal arrays and so forth. The name has been used for a special set of designs based on mutually orthogonal latin square designs, but it is not restricted to this category. In some experimental work, it is appropriate to deal with two sets of orthogonal arrays; (*1*) inner array or design array that includes those factors which may be deliberately set at appropriate levels by the experimenter and, (*2*) outer array or noise array that include those factors which ordinarily are not set at tightly specified levels. The levels of the outer array are set to reflect the expected level of noise.

orthogonal contrasts, *n*—two contrasts are orthogonal if the contrast coefficients of the two sets satisfy the condition that, when multiplied in corresponding pairs, the sum of the products is equal to zero. See **contrast** and **contrast analysis.**

DISCUSSION—*Example 1:*

	A_1	A_2	A_3
a_{i1} Contrast 1	−1	0	+1
a_{i2} Contrast 2	0	−1	+1
$a_{i1}a_{i2}$	0	0	+1

$\sum a_{i\,1}\ a_{i2} = 1 \therefore$ not orthogonal

Example 2:

	A_1	A_2	A_3
a_{i1} Contrast 1	−1	0	+1
a_{i2} Contrast 2	−1	+2	−1
$a_{i1}a_{i2}$	+1	0	−1

$\sum a_{i1}\ a_{i2} = 0 \therefore$ orthogonal

partially balanced incomplete block design (PBIB), *n*—an incomplete block design in which each block contains the same number k, of different versions from the t versions of the principal factor.

NOTE 13—The arrangement is such that not all pairs of versions occur together in the same number of the blocks; some versions can therefore be compared with greater precision than others.

DISCUSSION—The design implies that every version of the principal factor appears the same number of times r in the experiment.

Example: $t = 6,\ k = 4,\ b = 6,\ r = 4,\ n_1 = 1,\ n_2 = 4,\ \lambda_1 = 4,\ \lambda_2 = 2$

		Versions of principal factor			
Block	1	1	4	2	5
	2	2	5	3	6
	3	3	6	1	4
	4	4	1	5	2
	5	5	2	6	3
	6	6	3	4	1

In this design every version occurs $r = 4$ times and if we start with any version (for example, version 1), we find $n_1 = 1$ version (for example, version 4) that appears together with version 1 in $\lambda_1 = 4$ blocks and $n_2 = 4$ versions (numbers 2, 3, 5, and 6) that appear together with Version 1 in $\lambda_2 = 2$ blocks. These parameters, n_1, n_2, λ_1 and λ_2, are the same whatever the starting version may be.

partially nested experiment—a nested experiment in which several factors may be crossed as in factorial experiments and other factors nested within the crossed combinations.

NOTE 14—It is not unusual to find that experiments consist of both factorial and nested segments. See **nested experiment.**

Plackett-Burman designs, *n*—a set of screening designs using orthogonal arrays that permit evaluation of the linear effects of up to $n = t - 1$ factors in a study of t, treatment combinations.

DISCUSSION—Plackett-Burman designs were among the earliest sets of designs using orthogonal arrays for the purpose of screening many average effects with minimum experimentation. Each average effect is aliased with higher order terms when $n - 1$ factors are studied in n runs.

randomization, *n*—the procedure used to allot treatments at random to the experimental units so as to provide a higher degree of independence in the contributions of experimental error to estimates of treatment effects.

NOTE 15—An essential element in the design of experiments is to provide estimates of effects free from biases due to undetected assignable causes within the experimental space. Randomization is a process to minimize this risk. The operational procedure for assignment *at random* involves the use of random numbers or some similar method for assuring that each unit has an equal chance of being selected for each treatment.

randomized block design, *n*—a design in which the experiment space is subdivided into blocks of experimental units, the units within each block being more homogeneous than units in different blocks.

NOTE 16—In each block the treatments are allocated randomly to the experimental units within each block. Replication is obtained by the use of two or more blocks, depending on the precision desired, and a separate randomization is made in each block.

DISCUSSION—If the whole of the experimental material, area or time is not homogeneous, it may be possible to stratify the material into homogeneous groups or blocks. This approach is one of the methods for controlling the variability of experimental units. For the completely randomized design, no stratification of the experimental units is made. In the randomized block design the treatments are randomly allotted within each block, that is, the randomization is restricted.

Example: Four treatments *A, B, C* and *D* are assigned at random to the experimental units in each of three blocks.

Block	1	B	A	C	D
	2	C	B	D	A
	3	B	C	A	D

randomized block factorial design, *n*— a factorial experiment run in a randomized block design in which each block includes a complete set of factorial combinations.

residual error, *n*—the difference between the observed result and the predicted value (estimated treatment response); Observed Result minus Predicted Value.

DISCUSSION—For the purpose of this definition, the term *predicted value* is understood to be the estimated treatment response determined from the empirical model derived from the data of the experiment using the assumed model. Residual error includes experimental error and assignable sources of variation not taken into account by the model. A comparison of the *residual error* with the *experimental error* can be used to assess the validity of the assumed model since the *residual error* may include both *lack of fit* and *experimental error* components. The variance of the *residual error* is usually measured in an experiment by subtracting the pooled sum of squares for terms included in the assumed model from the total sum of squares and dividing by the corresponding difference in *degrees of freedom*. See experimental error, assumed model and regression analysis.

response surface, *n*—the pattern of predicted responses based on the empirical model derived from the experiment observations.

DISCUSSION—A sequential form of experimentation is often used in conjunction with the mapping of response surfaces in which the responses of the earlier stages are used to determine where to select additional treatment combinations for study so as to help find the optimum region efficiently. This approach is termed *response surface methodology.*

screening design, *n*—a balanced design, requiring relatively minimal amount of experimentation, to evaluate the lower order effects of a relatively large number of factors in terms of contributions to variability or in terms of estimates of parameters for a model.

NOTE 17—In screening designs, the term lower order effects is sometimes limited to first order terms such as linear components of main effects, but often includes both first order terms and second order terms such as two factor interactions and quadratic curvature components of main effects.

DISCUSSION—Screening designs are frequently used in sequential experiment programs when there are many potential factors contributing to variation. Typical screening designs are Plackett-Burman and other orthogonal arrays and small fractional factorial designs. Screening designs are particularly useful.

staggered nested experiment, *n*—a nested experiment in which the nested factors are run within only a subset of the versions of the first or succeeding factors.

DISCUSSION—In the example for a fully nested experiment, Version C_3 or C_4 and C_7 or C_8 might be eliminated, so that factor *C* is studied in only versions 1 and 3 of factor *B*. In this arrangement, the variability of *C* would be estimated with only half the precision of the arrangement for the fully nested experiment.

treatment, *n*—a combination of the levels (versions) of each of the factors assigned to an experimental unit. synonym **treatment combination**.

treatment combination, *n*—see **treatment**.

Youden square, *n*—A type of block design derived from certain Latin squares by deleting, or adding, rows (or columns) so that one block factor remains complete blocks and the second block factor constitutes balanced incomplete blocks.

DISCUSSION—*Example 1:* block factor 2 (columns)

block factor 1 (rows)	1	2	3	4	
1	A	D	C	B	
2	B	A	D	C	
3	C	B	A	D	
/4/	D/	/C/	B/	/A/	deleted from the Latin square

The elimination of the 4th row of the 4 × 4 Latin square yields this 3 × 4 Youden square.

Example 2: If the columns in the example under the balanced incomplete block design were considered as blocks (a second block factor), it will be seen that 3 columns from a 7 × 7 Latin square have been ignored, and the design would be a Youden square.

5. Keywords

5.1 experimental design; factorial experiment; statistics; terminology