Factorial Design in the Study of Chemical Reactions

HERBERT GROHSKOPF

AMERICAN CYANAMID CO., NEW YORK, N. Y.

The factorial approach accelerates and enhances experimental work when theoretical considerations do not suffice and a problem must be attacked frontally and with empirical methods. In a general discussion of the factorial technique, the evaluation of main effects and interactions is illustrated in this paper. The value of randomization of the order of experimental runs, and of the use of blocks, to improve experimental efficiency is discussed. Four independent variables in an organic synthesis in which a gas is reacted with a liquid were studied on a pilot plant scale. The factors were mole ratio of gas to liquid, retention time, reactor pressure and temperature, and the experiment was run as a standard factorial design. While this experiment was on a pilot plant scale, the use of factorial designs for the study of reaction velocities is proposed.

IN RECENT years there have been numerous articles proposing the use of the factorial design in experimental work. To date, published reports describing experimental work continue, for the most part, to be innocent of attempts to put the factorial method to work. The fact that much empirical work would be accelerated and enhanced by the factorial approach makes this lack the more regrettable. With the strong tradition of the one-factor approach to chemical experimentation on the one hand, and the diffuse literature of a new small-sample statistics, with its own special jargon on the other, this is no small wonder. In the long run the factorial method will come into broad use. Courses in the applications of small-sample statistics will become standard in the training of chemists and engineers.

In this nonmathematical discussion some of the implications of the factorial approach to experimental work are examined. Broadly speaking, factorial investigations fall into two classes:

Model I. Experiments in which the effects of controlled changes in the levels of several independent variables, or of several specific treatments, are to be observed and estimated in terms of shifts in the process mean. The example given in this discussion is of this type.

Model II. Experiments by which the possible causes of uncontrolled process fluctuation can be separated and evaluated. Examples of problems of this type are studies of variability of fibers, yarns, and filaments, variability between lots of a material, or between samples and analyses of the same material.

In Model I experiments the object is to determine significant changes in the process mean. In Model II experiments the object is to study the causes and magnitude of uncontrolled fluctuations about the process mean. Fortunately, the algebraic identities and assumptions which are basic to the analysis of variance are equally valid for Models I and II (*2*). The discussion in this paper is confined to consideration of a Model I experiment.

ORGANIZATION OF FACTORIAL EXPERIMENTS

Consider a system in which there are three major independent variables, or factors. Suppose, for the sake of simplicity, it is desired to measure the effects of a shift from the low level to the high level of each of the three variables on the process mean. Typical factors and levels might be those shown in Table I.

The possible combinations of three factors at two levels consist of 2^3 or 8 experimental runs. Fit into a balanced block, these runs assume a basic order (Figure 1). The subscript 1 designates the low level; the subscript 2 the high level of each independent variable, or factor. Within the low and the high level groupings of retention time, the other variables, M and T, are balanced as to level. One can, therefore, estimate the effect of raising the level of R from R_1 to R_2 as follows:

$$\text{Effect of } R = \text{runs } (3 + 4 + 7 + 8)/4 - \text{runs } (1 + 2 + 5 + 6)/4$$

The result is an estimate of the effect of the increase in the level of R, while the effects due to operation at two levels of M and T are averaged and, therefore, cancel out. The average effects of changes in levels of the other two variables are estimated by similar selective groupings. The average effect, observed with the other variables in their several combinations, is a more valid estimate of the effect of a change in the operating level of a variable than the point estimate which results when the other variables are held constant. It has been said that the experimental base of inference is broadened by the factorial method. It is also notable that each factorial run is made to contribute a result on the effect of operating at either the low or the high level of each independent variable.

	R_1		R_2	
	T_1	T_2	T_1	T_2
M_1	1	2	3	4
M_2	5	6	7	8

EIGHT—RUN BLOCK FOR A THREE FACTOR EXPERIMENT

Figure 1. A Factorial Replication

Effect of M = runs $(5 + 6 + 7 + 8)/4$ − runs $(1 + 2 + 3 + 4)/4$
Effect of R = runs $(3 + 4 + 7 + 8)/4$ − runs $(1 + 2 + 5 + 6)/4$

How the Factorial Experiment Defines a System. Roughly speaking, the factorial treats the chemical reaction system as an n-dimensional volume, much as in thermodynamics the properties of substances are described by curvilinear volumes, usually in three dimensions. When two independent variables are

Table I. Typical Factors and Their Levels

Independent Variable or Factor	Low Level	High Level
M = mole ratio of reactants (moles of A/mole of B)	5/1	8/1
R = retention time (average time of residence in reactor/particle of mass), min.	20	34
T = reactor temperature, ° C.	−5	+10

drawn as abscissa and ordinate, the per cent yield of a reaction system can be represented by constant yield lines, much as elevation contour lines are used in military mapping. The object of the experiment is to probe for peaks in yield at or near the most favorable conditions of operation. The choice of ranges of factor levels is guided by theoretical considerations and is limited by the design of the reactor. The pattern of a design on two factors is shown in Figure 2. The family of curves and typical results in Figure 2 were devised for illustrative purposes only.

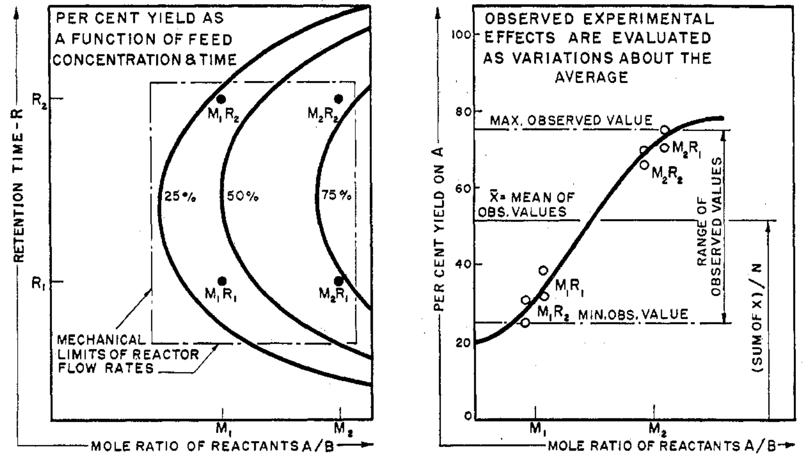


Figure 2. The Factorial Experiment

The factorial establishes a probing pattern by which the average effect of shifting each of *N* factors over a given range can be estimated

M_1 = Low mole ratio
M_2 = High mole ratio
R_1 = Low retention time
R_2 = High retention time

Two factors at two levels require 2^2 or 4 runs per replication:

$$\begin{matrix} M_1R_1 & M_2R_1 \\ M_1R_2 & M_2R_2 \end{matrix}$$

The profile of duplicate runs at these conditions gives an indication of the relative magnitudes of the factorial effects of M and R. The large vertical displacement between all runs at the low and high levels of M stands out as the preponderant effect. This factorial effect is an estimate of the response in the system to a shift from the low to the high level of M.

A smaller displacement of runs at low as opposed to high R is also apparent. From the point of view of process economics, M_2R_1 would appear to be the most favorable of the four operating conditions.

With a greater number of variables, the spatial simplicity of this example is lost. But the generalized case of any number of variables in an n-dimensional volume is no more complicated a problem in terms of factorial analysis than is the present case of two variables. Thus, in planning a factorial experiment, the chemist defines a set of minimum and maximum levels for each of n variables. For a given reactor, his choice of levels defines a segment of practical interest, which is selected from the total possible ranges of operating conditions for that reactor. Where desired, intermediate levels can, of course, be introduced with a corresponding increase in experimental requirements.

Aspects of Variation. A further contribution of the factorial design is its high experimental efficiency in the face of uncontrolled process variation. There are three fundamental types of uncontrollable variation in a continuous system:

1. Random Variation. An equal scatter of points below and above a normal or average operating level. Here the term operating level refers to measurements of dependent variables, and is defined as an average level of conversion, activity, or yield. This random variation about an average operating level usually falls in the familiar pattern of the normal distribution about a mean. This type of random variation is due to normal fluctuations in control, in flow rate and flow composition, and to errors of sampling and determination.

The second and third types of variation can be of serious consequence to the experimenter. They result from the fact that in a complex system other things almost never remain the same:

2. Shifts in Level. Given a level of operation for an extended period of time, it is not unusual to find a sudden rise or fall in the system to a new steady-state level of operation. While physical and chemical causes may be operating to bring such changes about, it is often impossible to prevent, to measure, or to control them. It is here that the nonrandomized one-factor experiment tends to fall down.

3. Trends in Level. Finally, a third kind of variation may be superimposed on the basic patterns of 1 and 2, above. In addition to normal variation and to intermittent jumps in level, chemical systems are not unfrequently subject to sustained linear or higher-order trends in level of operation. These may be associated with uncontrollable changes in the system with time, such as a gradual build-up of impurities, increases or decreases in catalyst activity, and so forth.

How Increased Experimental Sensitivity Is Realized. The factorial method relies on two devices to cope with uncontrolled variation:

1. Randomization. The actual order of runs in a factorial experiment is chosen by chance. In a two-level experiment, for instance, half the runs are at the low level and half are at the high level of each factor. It follows that the chances that one, and only one, level of a variable might be associated with sudden uncontrolled shifts in the operating level of the system is minimized. A similar process of random distribution of runs is carried out in factorials run at more than two levels. The importance of random order to an unbiased estimation of treatment effects and of the experimental error cannot be overemphasized.

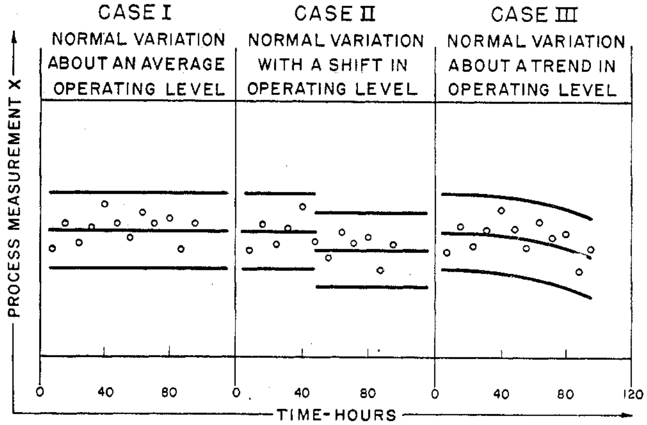


Figure 3. Patterns of Process Variation

2. Small Experimental Blocks. Whenever experimental control over extended periods of time presents a problem, it is possible to subdivide the factorial experiment into balanced blocks of small run size. A block of four runs, for instance, can be run in random order during four consecutive shifts (Figure 4).

Assume that a given process is known to be subject to considerable variation with time. If one could estimate the effects of changes in the levels of the independent variables from the smaller, within-block time base, the precision of determination would be improved as compared with a measurement of experimental effects from the average value of the larger number of runs for the entire experiment.

The utility of small block designs becomes apparent from the example shown in Figure 5.

The dependent variable reported in this case was the per cent of formation of an undesired gas addition product. The line marked $\overline{X}$ represents the grand experimental average of 32 runs, and only half of a total of eight four-run blocks are charted. Since within-block levels of the independent variables are balanced, the effects of changes in the level of a factor are estimated as deviations from the block mean. In other words, the within-block comparisons are made in such a manner that they are free of between-block variation.

	M_1		M_2	
	R_1	R_2	R_1	R_2
T_1		3	2	
T_2	4			1

A BALANCED FOUR—RUN BLOCK FOR A THREE FACTOR EXPERIMENT

Figure 4. Four-Run Block

High level of a variable is positive in sign
Low level is negative
***MR* interaction is determined by algebraic multiplication of signs:**

$$\overset{M_2R_2}{[(+)(+)(\text{run }1)]}\overset{+}{+}\overset{M_1R_1}{[(-)(-)(\text{run }4)]}\overset{-}{+}$$
$$\overset{M_2R_1}{[(+)(-)(\text{run }2)]}\overset{-}{+}\overset{M_1R_2}{[(-)(+)(\text{run }3)]}$$

Where there is no appreciable process variation associated with time, greatest experimental sensitivity is achieved when all the runs in a factorial replication are carried out in one randomized block. In this case the experimental effects are estimated from the mean operating level of the entire randomized block or replication.

EXPERIMENTAL

An experiment which was carried out at the Stamford Laboratories of the American Cyanamid Co. provides an example of the use of these techniques. The engineers who carried out the experiment in 32 experimental runs made a comparison of the efficiencies of the factorial and the one-factor approaches. They estimate that the information gained from these 32 runs could have been duplicated by carrying out over 100 such runs with the one-factor method of experimenting. This estimate appears to be conservative. It has been stated that the information gained per unit of experimental work can be increased six to ten times by a well-organized factorial (*3*).

Several considerations make it impossible to give exact details of the process at this time. It may be of interest that it was carried out in a reactor large enough to process over 120 pounds of raw materials per hour.

A gaseous feed stream was reacted with a feed stream consisting of an organic liquid. The reactor was of a single-stage type, and the reaction was catalyzed by a sulfuric acid catalyst. Refrigeration was applied to the reactor to repress the rate of sulfonation. The gas was readily absorbed in the liquid phase, and good mixing was assured by circulation velocities in the turbulent flow range. The product stream was neutralized in a caustic soda wash and was purified by distillation after a separation of the product layer.

Table II. Independent Experimental Variables

T	= Reactor temperature, ° C.	−5-+10
P	= Reactor pressure, lb./sq. in. gage	25-30
R	= Retention time (average time of residence of mass in the reactor), min.	20-34
M	= Reactant mole ratio (moles of gas/mole of liquid in the feed stream)	5/1-8/1

The factors and their levels are shown in Table II. Four factors at two levels call for 2^4 or 16 runs per replication for a balance of factors at two levels. A set of all possible combinations of each factor at *n* levels is called a replication. A replication consists of 2^3 or 8 runs for three factors, and of 2^4 or 16 runs for four factors at two levels.

Note that the two replications were grouped in four blocks each. The key to the factorial symbolism is as follows: The presence of a letter indicates that the factor it represents is to be held at the high level for a given run. *MR* designates a run with high mole ratio and high retention time, with the other two factors at their low level. *TPMR* is a run with all factors at their high level and [1] is a run with all factors at their low level (*1*). The meaning of contrasts which are confounded will be treated in the discussion of interactions.

Table III. Experimental Design

Replication 1—*TP*, *TMR*, *PMR* Confounded				Replication 2—*MR*, *TPM*, *TPR* Confounded			
Blocks				Blocks			
1	2	3	4	5	6	7	8
[1]	*TP*	*T*	*P*	[1]	*MR*	*M*	*R*
TPM	*M*	*PM*	*TM*	*TMR*	*T*	*TR*	*TM*
TPR	*R*	*PR*	*TR*	*PMR*	*P*	*PR*	*PM*
MR	*TPMR*	*TMR*	*PMR*	*TP*	*TPMR*	*TPM*	*TPR*

Process Measurements. The following process measurements were reported and analyzed:

Conversion of liquid feed
  Total liquid feed less unreacted liquid withdrawn in product stream
Yield on liquid
Yield on gas
Per cent residue
  Consisting of high boiling reaction products of the gas

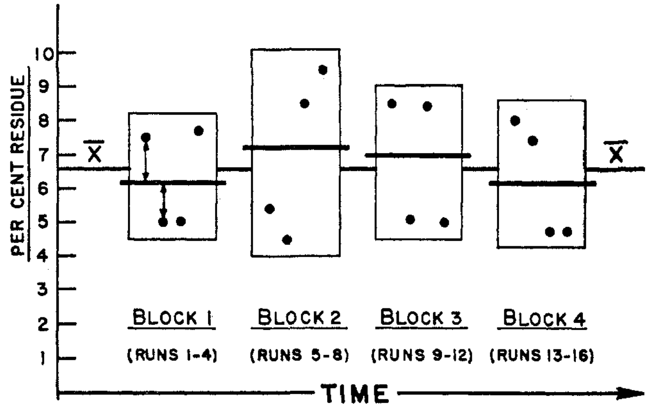


Figure 5. Blocks 1-4 on Per Cent Residue

$\overline{X}$ is the grand average of eight blocks

The effect of changing from the low to the high level of each of the independent variables is shown in the table of factorial effects (Table IV). Mole ratio was the preponderant main effect, and indeed the only one to attain high statistical significance. Retention time was just significant on gas yield and on per cent residue. But in addition to the evaluation of main effects due to independent variables, the factorial makes a routine screening of effects which are associated with more than one factor.

Interactions between Factors. In order to illustrate the interaction between two or more factors, the grouping of results on per cent yield based on the gas feed is made as shown in Table V. Each number represents the average of eight experimental trials.

Table IV. Averages and Factorial Effects

Factor	$\bar{x}$ Factorial Effect Conversion of liquid feed	Yield on gas feed	Residue, %
M = 5	30.73	59.73	8.36
M = 8	19.27 (−11.26)	66.65 (+6.92)	4.85 (−3.51)
P = 25	25.24	62.56	6.70
P = 30	24.98 (−0.25)	63.81 (+1.25	6.51 (−0.19)
R = 20	25.32	63.90	6.76
R = 34	24.88 (−0.44)	62.48 (−1.42)	6.46 (−0.30)
T = −5	25.34	63.71	6.56
T = +10	24.98 (−0.48)	62.67 (−1.04	6.64 (+0.08)

Table V. *MR* Interaction—Yield Based on Gas

(Each value is average of four runs)

	R = 20 Min.	R = 34 Min.
M = 5/1	59.25	60.80
M = 8/1	69.95	63.35
Difference	10.70	2.55

The classification is made according to levels of M and R. In graphical terms the relationship illustrated in Figure 6 was observed. Mole ratio was highly significant as a main effect (1% level). Retention time was just significant as a main effect (5% level). The graph of the highly significant interaction shows that yield rose with an increase in retention time at the low mole ratio, and dropped with an equal increase in retention time at the high mole ratio of gas to liquid feeds.

An interaction can be defined as a dependence of the effect on process performance of an increase in variable A on the simultaneous level of variable B. In this case the increase in yield due to a stated increase in mole ratio is dependent on the length of retention time. Interactions are quite common in chemical reactions.

As will be seen in later discussion, the sum of squares generated by each potential source of variation is calculated in the analysis of variance. For the *MR* interaction the sum of squares on the gas yield is 68.06. Interesting relationships can be examined by noting that the sum of squares estimated for the *MR* interaction on per cent conversion is 0.81, and on per cent residue, 0.00. For these other measures of process performance no *MR* interaction was observed. The constant mole ratio lines describing the *MR* interaction on these results did not depart from the parallel. This is simply another way of saying that there was no interaction.

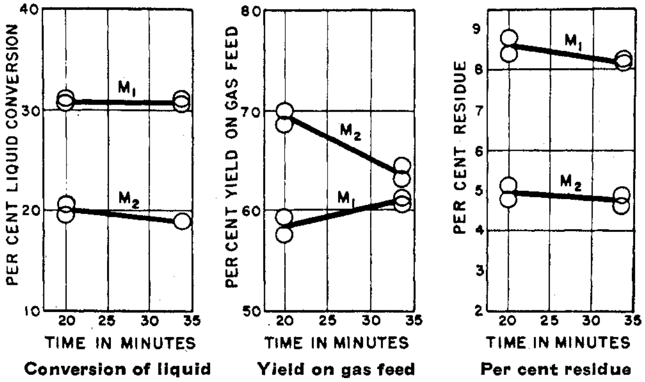
Other significant interactions were the *MP* and the *PR* interactions on per cent residue, and the *MTR* interaction on per cent conversion. At this point a belated note on the details of the design of the present experiment may be considered. In each replication one first-order and two second-order interactions are confounded. This means that the efficient small-block organization was achieved at the price of foregoing each of these comparisons in one of the replications. If a positive sign for the high level of the variable and a negative sign for the low level is assumed, cross-multiplication of signs gives the level of the interaction in Table VI.

Table VI. *MR* Interaction

Replication 1—*TP*, *TMR*, *PMR* Confounded Block 1	*M*	*R*	*MR*	Replication 2—*MR*, *TPM*, *TPR* Confounded Block 4	*M*	*R*	*MR*
[1]	−	−	+	*R*	−	+	−
TPM	+	−	−	*TM*	+	−	−
TPR	−	+	−	*PM*	+	−	−
MR	+	+	+	*TPR*	−	+	−

It is clear that a within-block comparison of *MR* can be made in replication 1. In replication 2 there is no way of separating the within-block level of the system, as biased by possible time trends, and the actual effect of operating at the low level of the *MR* interaction. The statistician says that the interaction and the block have been confounded.

All potential interactions between the factors are screened for significance in a two-level factorial with a minimum of experimental effort. The criticism that a probe at two levels does not define a curve is valid. On the other hand, where important interactions exist and where reasonable care has been taken in the choice of levels, the factorial will reveal their existence. An interaction such as the *MR* interaction is an important result to the process engineer. Without a measure of its existence he would be seriously handicapped in design work. To the physical chemist, however, an interaction is a question rather than an answer. Sometimes this question may fall apart rapidly in the light of known physical and chemical relationships; sometimes it represents a point of departure for further investigations. The kineticist may find that he can formulate the statistically estimated interaction in the more rigorous terms of differential equations. Thus the statistically estimated interactions, though not definitive in a theoretical sense, is a durable measure of the existence of physical and chemical relationships. It is durable in the sense that it is not likely to be lost sight of in the research-development history of the process. It is inherently ordered, in that the screening of the system for possible interactions is usually complete. It is definitive, for while it may not explain a set of relationships, it reveals controlling factors in that relationship.



Conversion of liquid   Yield on gas feed   Per cent residue

Figure 6. *MR* Interaction

Moles of gas feed per mole of liquid feed

M_1 = 5/1
M_2 = 8/1

ANALYSIS OF VARIANCE

A measure of the total observed variation for the entire experiment is obtained when the individual observations are squared and summed, and the sum so obtained is corrected to make it independent of the average experimental level:

$$\text{Total sum of squares} = (\text{sum of } X^2) - (\text{sum of } X)^2/n$$

Where the total sum of squares is a measure of the total observed variation about the grand average of all observations, X = the individual observation, and n = total number of observations.

The term (sum of X)$^2/n$ is called the correction factor. Subtracting it from the (sum of X^2) removes the sum of squares generated by the average observed level of X. What is left is an estimate of a band of variation which falls equally above and below the average level of X for the experiment. Reference to Figure 2 serves to illustrate this concept.

In the analysis of variance each mean square is a measure of the variability associated with each treatment. An estimate of

Table VII. Mean Squares Obtained by Analysis of Variance

Source of Variance	D.F.	Conversion of Liquid Feed, %	Yield on Liquid Feed	Yield on Gas Feed	Residue, %
Replications	1	0.21	9.46	0.45	0.05
Blocks within replications[a, b]	6	2.79	30.58	21.58**(r)	0.47*(r)
$M^{a,b}$	1	1,014.75***(r)	5.61	383.64***(r)	98.35***(r)
$P^{a,c}$	1	0.50	18.30	12.50*(p)	0.30
T	1	1.81	3.25	8.61	0.05
$R^{a,b,c}$	1	1.53	7.03	15.96*(p)	0.69*(r)
MP	1	0.05	1.20	1.05	1.09
MT	1	0.25	4.06	3.13	0.03
$MR^{a,b}$	1	0.81	49.70	68.06**(r)	0.00
PT	1	0.12	1.32	0.49	0.18
$PR^{a,b}$	1	2.88	7.03	1.62	0.75*(r)
TR	1	3.92	10.81	0.55	0.01
MPT	1	7.56	8.41	5.76	0.36
$MTR^{a,b}$	1	14.44*(r)	33.06	0.01	0.05
MPR	1	0.00	2.02	0.64	0.23
TPR	1	0.42	54.02	10.89	0.01
$TPRM^{a,b}$	1	10.81*(r)	94.53	13.78	0.03
Residual	9	1.36	16.40	3.50	0.12
Pooled error estimate		0.94	15.17	2.51	0.11
Pooled D.F.		19	27	18	19

[a] Significance level: *5%, **1%, ***0.1%.
[b] (r) signifies a test against the residual.
[c] (p) signifies a test against the pooled error estimate.

the effect of operating at two levels of mole ratio within the structure of the experiment is obtained by summing the values of yield on the gas observed at the high level and low level runs:

$$\begin{array}{lcr} \text{Sum of } M_2 - \text{(high level)} & = & 1066.6 \\ \text{Sum of } M_1 - \text{(low level)} & = & \underline{955.8} \\ \quad \text{Difference} & = & 110.8 \end{array}$$

$$\text{Average effect} = 110.8/16 = 6.9\% \text{ on gas yield}$$

The sum of squares due to this factor is simply the difference squared and divided by the total number of runs:

$$(110.8)^2/32 = 383.65$$

Since the total sum of squares on the gas yield was found to be 688.14, it is found that the variation due to operation at two levels of M accounted for over 55% of the total observed variation in the system. Other mean squares are calculated by the same or by similar methods until all known sources of variations have been assigned an estimated portion of the total variation (Table VII).

Error Estimates. Operating conditions such as level of temperature, mole ratio, and retention time are fairly reproducible, and are said to be members of a finite population of such conditions. They are Model I variables, and their application is subject to experimental control. Factors of this type must be distinguished from Model II factors such as raw material lots, days, operators, and tensile test pieces, which are members of infinite populations. Clearly, two raw material lots are just two representatives of the total population of all possible raw material lots which might be processed. Factors which are members of infinite populations tend to introduce more variability in an experiment than do factors from a finite population.

The procedure for estimating error for Model I experiments will now be described (*2*). When the component sums of squares due to temperature, pressure, mole ratio, retention time, their interactions, and all other identifiable sources of variance have been found, they are summed and subtracted from the total sum of squares. The remainder is called the residual. The residual is a first approximation to an error estimate.

The test of significance in the analysis of variance is simple. The size of the ratio

$$F = \frac{\text{estimate of mean square due to an experimental effect}}{\text{estimate of error mean square}} = \frac{\sigma^2 \text{ effect}}{\sigma^2 \text{ error}}$$

is a measure of the probability that an effect is real. If the ratio is large, the effect is probably real. If it is small, the effect is probably nonexistent. Tables for the value of the F (or variance) ratio give the size the F ratio must equal or exceed at specified levels of probability in order that the effect may be considered significant. A ratio significant at the 5% level of F implies that so high a value will be attained or exceeded in only about one out of twenty experiments where there is no real effect.

Degrees of Freedom. Each comparison in the analysis of variance demands one or more associated degrees of freedom. To make a comparison between the low and the high level of a factor demands one degree of freedom. The estimation of between-block variation demands n-1 or three degrees of freedom for each replication, since there are four blocks in each replication. The total number of degrees of freedom associated with an experiment is equal to the total number of observations made less the number of parameters to be estimated. It will be seen that degrees of freedom assigned to nonsignificant comparisons can be pooled with the residual error along with their sums of squares in order to obtain a more sensitive estimate of error.

Pooling of Error Estimates. In order to make full use of the information inherent in the results of a factorially designed experiment the following simple rule can be applied to Model I experiments: "Pool with the error estimate all mean squares not greater than twice the original error mean square" (*4*, *5*).

Mean squares which are tested and found to equal or exceed an F ratio of 2.0 are set aside and do not enter in the pooling. A more sensitive error estimate is thus obtained by pooling sums of squares and degrees of freedom where this criterion is met. Estimates pooled in accordance with this rule appear at the bottom of the analysis of variance (Table VI).

As an example, the pooling of the sums of squares on the yield based on gas feed is shown in Table VIII.

Table VIII. Pooling of Error

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares
Residue	9	31.52	3.50
Replications	1	0.45	
MPR	1	0.64	
MTR	1	0.01	
MPT	1	5.76	
TR	1	0.55	
PR	1	1.62	
PT	1	0.49	
MT	1	3.13	
MP	1	1.05	
Pooled error estimate	18	45.22	2.51

TESTS OF SIGNIFICANCE

The test for significance of between-block variation on the gas yield data follows as an example. The degrees of freedom (D.F.) associated with blocks can be grouped as follows:

$$(4 - 1) \text{ D.F./replication} \times (2) \text{ replications} = 6 \text{ degrees of freedom}$$

The sum of squares due to blocks = 129.50, and

$$F = \text{mean square ratio} = \frac{129.50/6}{45.22/18}$$

$$= 21.58/2.51. = 8.60^{***}$$

The significant mean square ratio with six and eighteen degrees of freedom must equal or exceed the following values:

Level of Significance	Significant Ratio	Notation
5%	2.66	*
1%	4.01	**
0.1%	6.3	***

Thus the between-block variation of the yield on gas feed proved significant at the 0.1% level. This can be taken to mean that the probability of observing so large a mean square ratio due to block-to-block variation in the absence of actual block-to-block variation is lower than about one chance in 1000. Where the sums of squares due to replications, blocks, main effects, and interactions are found to be significant, they are set aside and do not enter into further calculations in the routine analysis of variance.

This brief excursion into the mechanical aspects of the analysis of variance leaves many aspects of the subject unmentioned, glossed over, or undefined. A number of books currently in print treat the subject in varying detail. There is, of course, no substitute for freely available statistical advice by the expert.

CONFIDENCE LIMITS

The engineer often values a yardstick as to the probable limits of a given estimate of performance. Such an estimate can be made once a valid estimate of error is available. To determine confidence limits for an estimated level of performance "Student's ratio," defined as:

$$t = \frac{\text{error in an estimated mean}}{\text{estimated error of the mean}}$$

is used. Tables of "Student's ratio" at several levels of significance are reproduced in most modern statistics books.

To calculate confidence limits on the mean value of yield associated with a given set of operating conditions the following expression is used:

$$\bar{x} \pm t_{0.05}(s_{\bar{x}})$$

where

$\bar{x}$ = the average observed yield at a given set of operating conditions based on two runs for each mean

$t_{0.05}$ = the t = ratio at the 5% level with 18 degrees of freedom = 2.10

$(s_{\bar{x}})$ = the estimated standard error, which is the square root of the gas feed error mean square divided by two runs per mean, or $\sqrt{2.51/2} = 1.12\%$

then the confidence limits = 2.10 (1.12%) = 2.35%

and the limits on the mean = $(\bar{x} \pm 2.35)$ %

Thus, the true mean yield for each set of operating conditions will lie within ±2.35% of the mean observed for the two runs at each operating condition in 19 out of 20 cases.

A similar comparison for confidence limits on the per cent residue with nineteen degrees of error freedom would be:

$$\text{Confidence limits} = \pm 2.09\ (0.235) = \pm 0.49\%$$

Thus, the true mean level of residue associated with any given set of operating conditions can be expected to fall within ±0.5% of the observed mean.

Significance of a Factorial Effect. To test for the significance of a factorial effect the following expression is used:

$$\left(\begin{array}{l}\text{Least significant difference}\\ \text{between treatment means}\end{array}\right)_{0.05} = t_{0.05}\sqrt{\frac{\text{Mean square} \times 2}{n}}$$

where n = the number of runs in a mean.

As an example, the factorial effect due to retention time on the gas yield is based on the difference between two means, each of which was obtained from 16 runs:

	% Yield
Average for 20-minute retention	= 63.90
Average for 34-minute retention	= 62.48
Factorial effect	= 1.42

Is this effect significant by the least significant difference (L.S.D.) test?

$$\text{L.S.D.}_{0.05} = 2.10\sqrt{\frac{2.51 \times 2\%}{16}}\ \text{in yield}$$

$$= 2.10\ (0.56)\% = 1.19\%\ \text{in yield}$$

This is less than the observed factorial effect of 1.42%, which is therefore significant by the test. This means that an effect larger than 1.19% would be obtained in only about one out of twenty trials if there were no real difference due to the change in operating level tested. The effect just misses falling within 2% L.S.D. limits of 1.43% on the yield.

SUMMARY

The preponderant main effect was that of mole ratio of the feed gas to liquid feed.

Temperature was found to contribute no significant main effects and comparatively unimportant interactions within the experimental ranges screened. Temperature can, therefore, be maintained at the higher level with a considerable saving in refrigeration capacity requirements and operating costs.

It was observed that the partial pressure of the gas and the time of retention interact, and that at the higher partial pressure, or mole ratio, the lower retention time is favored. This finding would have considerable influence on future design work on this process.

When to Use the Factorial. It is apparent that this experimental tool will answer only those questions which the organization of the experiment poses. A chemist will approach the problem of organizing a factorial design differently than an engineer. In rate studies, for instance, such a variable as mole ratio might be recast in its purer, and less convenient, form as concentration.

It has been said that the ideal stage in the research history of a process at which to apply the factorial is in the first experimental probing of the system. Investigators are frequently in a state of relative ignorance concerning the relationships of a number of variables. The factorial is ideally suited to afford a rapid ordering of notions about a system. A half-replicate of six factors, for instance, requires 32 runs. It permits sixteen observations at a low level and sixteen observations at a high level of each of six variables. It screens the fifteen possible two-factor interactions, although the higher-order interactions are lost in this design.

A factorial should be run to characterize and formulate the findings of chemical research and of development work on each project. As a source of engineering design information it is comprehensive. Conclusions are based on operation with a diverse set of conditions. It has been said that the factorial broadens the base of experimental inference. The relationships between variables are clear from the tables of main effects and the analysis of variance. The associated tests of significance and estimates of error are invaluable for a realistic gaging of process economics and design requirements.

Perhaps it is a tendency of chemists and engineers to isolate physical and chemical phenomena and to study them as things apart. This tendency has been persistent in many fields of science. In the process, the description of natural systems operating in nonideal environments has been slighted. The present thesis, therefore, is that while the study of isolated and idealized systems is essential to basic progress, much of nature can be known only from the study of the complex of factors which are at work in a continuous system at steady state. It is in this direction that the factorial can make a considerable contribution.

ACKNOWLEDGMENT

The help and cooperation of project engineers, H. P. Callahan and H. Wallman, are acknowledged. F. Wilcoxon, company statistician of American Cyanamid Co., made much-needed statistical advice available. E. Baumgarten and H. M. Hulburt, American Cyanamid chemists, and J. C. Whitwell of Princeton

University were helpful in discussing individual aspects of the subject matter. Experimental studies were carried out at American Cyanamid Co. Research Laboratories, Stamford, Conn.

LITERATURE CITED

(1) Cochran, W. G., and Cox, G. N., "Experimental Designs," pp. 154–69, 209, New York, John Wiley & Sons, 1950.
(2) Eisenhart, C., *Biometrics*, 3, No. 1, 1–21 (1947).
(3) Mueller, R. K., "Effective Management through Probability Controls," pp. 132, New York, Funk & Wagnalls Co., 1950.
(4) Paull, A. E., "On a Preliminary Test for Pooling Mean Squares in The Analysis of Variance," *Ann. Mathematical Statistics*, **XXI**, 539 (1950).
(5) Tukey, J. W., presented at the Middle Atlantic Conference, Am. Soc. Quality Control, New York, 1952.

RECEIVED for review February 10, 1953. ACCEPTED April 20, 1953.

Phenol Synthesis by Direct Oxidation of Benzene

JU CHIN CHU, H. C. AI, AND DONALD F. OTHMER
POLYTECHNIC INSTITUTE OF BROOKLYN, BROOKLYN 1, N. Y.

The work was undertaken to explore a new method of direct oxidation of benzene to phenol by electric discharge, free from the difficulties usually associated with other direct oxidation methods, such as the presence of secondary reactions at high temperatures and preferential adsorption of phenol on catalysts. Within the limitations of the initial design of the apparatus, the optimum operating pressure was found to be 6 to 12 mm. of mercury at 3000 to 4000 volts of the discharging circuit, giving a conversion rate of 10 to 12.5 mole % of phenol per pass at a benzene throughput of about 0.1 gram per minute. The temperature effect was found to be unimportant in the range of 100° to 340° C. investigated. The experimental data have been correlated by the expression: $y = ABz + C$, where y = phenol yield per pass in 10^{-5} gram-mole per minute, $A = 1.670 + (x - 1)^{-0.21} + 0.185\,(x - 9)^{-4}$, $B = 2.800\,(0.879)^{1.35z} + 3.834\,(0.726)^{1.10x} - 2.135$, $C = 0.072\,(z - 0.4)^{-1}\,(x - 6)^{-2} - 0.620\,(z - 0.4)(x - 1)^{-4}$, z = benzene pressure in millimeters of mercury and x = oxygen pressure in millimeters of mercury. Equipment efficiency may be improved by providing a more compact discharge chamber and a larger surface of contact for the reactants. The yield per pass can be expected to increase materially for possible commercial application.

THE present commercial processes for producing phenol are all indirect, requiring the preparation, isolation, and purification of intermediate compounds such as benzene sulfonate, chlorobenzene, and cumene. The handling and consumption of critical materials such as chlorine, sulfuric and hydrochloric acids, sodium hydroxide, and propylene have added another disadvantage. The obvious advantage of a direct oxidation process has encouraged many investigations in the past, but so far the yields have been low.

A summary of all available pertinent data for the aforementioned direct oxidation processes is to be found in Table I.

For the reaction

$$C_6H_6(g) + {}^1/_2O_2(g) = C_6H_5OH(g)$$

the final standard free energy and equilibrium constant equations are found (*1*) to be

$$\Delta F_T = -41{,}728 + 17.88T - 1.082T - 6.45 \times 10^{-4}T^2 - 1.83 \times 10^{-8}T^3 \quad (1)$$

$$\log K_T = -\Delta F_T/4.576T \quad (2)$$

These two equations indicate that practically quantitative equilibrium conversions can be obtained under all ordinary operating temperatures and pressures. In other words, the thermodynamics of the oxidation reaction imposes no restrictions to process conditions.

KINETICS OF THE OXIDATION REACTION

The structure of benzene is a very stable configuration with respect to thermal or chemical attack. According to Jost (*27*), the C—H bond energy of benzene is around 102 kcal. per mole and the dissociation constant is only 0.0034 even at 700° C. Amiel (*2*) and Hinshelwood and Fort (*26*) found the apparent heat of activation for benzene oxidation to be the second highest among all the hydrocarbons (methane being the highest), 50 to 56 kcal. per mole. As the dissociation of the first single C—H bond is realized only at fairly high temperatures, and such dissociation must precede any chemical reaction, the primary oxidation product, once formed at such high temperature conditions, is usually less stable than benzene itself and, consequently, is subject to further attack by the reactant. Also, the dissociated phenyl radical may dimerize at such conditions. Even though the monosubstituted compound may form in quantity, such substituted functional group may often serve to activate other positions on the benzene nucleus, resulting in polysubstituted byproducts.

In order to moderate the oxidation reaction so that it may not proceed too far beyond the phenol stage, the following means could be employed:

1. Using a limited supply of oxygen, or diluting the oxygen with inerts.
2. Providing means of rapid heat removal from the reaction zone.
3. Using a temperature as low as possible.
4. Using a space velocity as high as possible.
5. Providing some means, such as a catalyst, a promoter, an electric discharge, or an exposure to actinic light, to activate the initial reversible reactions without exposing the reaction mixture unduly long to high temperatures.

At low temperatures and low oxygen concentrations, with high space velocity or short residence time in the reaction zone, the oxidation reactions of higher order could be retarded; but, in the absence of the agents as mentioned in paragraph 5 above, the yield per pass of the primary oxidation product could also be reduced to insignificant proportions. It is, therefore, obvious that a combination of the methods mentioned in 2 and 5 would give the best results. The right kind of activating agent remains to be chosen.

The failure of Weiss and Downs (*45*) and Charlot (*9*) to obtain any significant yield of phenol by catalytic oxidation of benzene in spite of their exhaustive search for a suitable catalyst has been attributed to the preferential adsorption of phenol on the catalyst.