Transactions of the Nigerian Association of Mathematical Physics
Volume 5,(September and November, 2017), pp241 –250

A Note on the Analysis of 3^n Factorial Experiments

Dawodu G.A[1], Apantaku F.S[1], Adeogun A.I[2] and Asiribo O.E[1]

[1]Department of Statistics, Federal University of Agriculture, Abeokuta Nigeria.
[2]Department of Chemistry, Federal University of Agriculture, Abeokuta Nigeria.

Abstract

Whenever a factorial experiment is not appropriately planned and its lay-out or randomization process is wrongly chosen, the analyses tend to be difficult and its interpretation is rendered meaningless and it becomes grossly inadequate. The 3^2 factorial, for example, is the easiest and simplest in the family of 3^n, $n \geq 2$ factorial experiments, yet if it is ill-planned with inappropriate lay-outs and randomization processes, it can be quite difficult to carry-out with clumsy and inexplicable results, for some research works. This paper purports to present valid and generalizable techniques from 3^n, n = 2, 3 factorials to higher members of the 3^n family in a simplified and reusable form. At least, an illustration of each proposed technique is also included for the readers' perusal and appreciation.

Keyword: 3^n series, experimental lay-out, randomization process, Yates' notations and Confounding.

1. INTRODUTION

In a factorial experiment the effects of a number of different factors are investigated simultaneously. The treatments (T_i, i = 1, 2, …, k) consist of all combinations that can be formed from the different factors [1, 2]. In this work, letters A, B, … shall denote factors whilst a, b, … will generally be for specifying levels of the corresponding factor, for the purpose of clarity; if a factor (A) has three levels then the symbols a_0, a_1, a_2 will denote respectively levels 1, 2 and 3. The testing of three variations of each factor makes possible a more thorough evaluation of the effects of the factors. With factorial experiments of this type, the F-test of the complete treatments means square (m. s.) may not be adequate at all times because it is directed, as it were, at a mixture of several diverse questions [3]. By an extension of the analysis of variance, we can subdivide the treatments sums of squares (s. s.) into a number of components that are more relevant to the individual questions. Moreover, an F-test can be made on the mean square for each component. The rules for subdivision are given in [2]. But for the purpose of this work, the subdivision into single components is explained thus:

Subdivision into single components: Let us first consider the simplest case in which all the treatments have the same number of replications. If we denote a treatment total as T_i, i = 1,2,…,n, where n is the number of treatments altogether. Then

$$T_1 - T_2, \quad n = 2, \qquad (1)$$
$$\frac{T_1 + T_2}{2} - T_3, \quad n = 3$$
$$T_1 + T_2 - T_3 - T_4, \quad n = 4 \quad ...$$

are some of the quantities (linear functions of the T_i's) that we are interested in. Since the quantities are comparisons amongst the T_i's, the sum of their coefficient is necessarily equal to zero. Consequently, for any comparison (linear function);

$$z_j = \alpha_{j1}T_1 + \alpha_{j2}T_2 + ... + \alpha_{jn}T_n \qquad (2)$$

1. $\alpha_{j1} + \alpha_{j2} + ... + \alpha_{jn} = 0$ and with r, as replication; $\frac{z_j^2}{D_j}$, $D_j = r\left(\alpha_{j1}^2 + \alpha_{j2}^2 + ... + \alpha_{jn}^2\right)$ will be the component of the sum of squares for treatments and will represent 1degree of freedom (d.f.) in the appropriate Analysis of Variance (ANOVA) table.
2. Any pair of comparisons, z_1 and z_2 are orthogonal if $\alpha_{11}\alpha_{21} + \alpha_{12}\alpha_{22} + ... + \alpha_{1n}\alpha_{2n} = 0$.
3. If z_1 and z_2 are orthogonal then $\frac{z_2^2}{D_2}$ is a component of (Treatment s. s - $\frac{z_1^2}{D_1}$).

Correspondence Author: Dawodu G.A., E-mail: abayomidawodu@yahoo.co.uk, Tel: +2348033753735)

4. If $z_1, z_2, \ldots, z_{k-1}$ are mutually orthogonal, then
   Treatment s. s = $\frac{z_1^2}{D_1} + \frac{z_2^2}{D_2} + \ldots + \frac{z_{k-1}^2}{D_{k-1}}$ (3).

The table below shows the single components most frequently used for interpretative purposes when there are three treatments. It is assumed that all treatments have the same number of replicates. The divisor D are those required for inserting the square of z in the analysis of variance.

Table 1: Sample sets of single components (Three treatments)

i) Equally spaced increments of one ingredient

Trt.	Z_1	Z_2
T_1	-1	1
T_2	0	-2
T_3	1	1
Component	Linear	Quadratic
Divisor D	2r	6r

ii) Two qualities of an ingredient and a control

Trt.	Z_1	Z_2
T_1 (0)	-2	0
T_2 (a_1)	1	-1
T_3 (a_2)	1	1
components	Effects of ‘a’	Quality difference
Divisor D	6r	2r

When different treatments have different numbers of replicates, then the following ‘measures’ are adopted:
If the i^th^ treatments has r_i replicates, then the treatments s.s. becomes;

$$\frac{T_1^2}{r_1} + \frac{T_2^2}{r_2} + \ldots + \frac{T_j^2}{r_j} - \frac{(T_1 + T_2 + \ldots + T_j)^2}{\sum_{i=1}^{j} r_i} \quad (4)$$

Also;
i) Equation (2) becomes a comparison amongst the treatment totals T_i if;
$\alpha_{j1} + \alpha_{j2} + \ldots + \alpha_{jn} = 0$
ii) Two components z_1 and z_2 are orthogonal if; $\alpha_{11}\alpha_{21} + \alpha_{12}\alpha_{22} + \ldots + \alpha_{1k}\alpha_{2k} = 0$

2. 3^2 Factorial Experiment

Let the factors be denoted by N and P. Consequently, the three levels of each of them are (n_0, n_1, n_3) and (p_0, p_1, p_2) respectively. This will enable us to capture the raw data (e.g. total yields over an equal number of plots, say, m), as contained in Table 2;

Table 2: The table of total yields (over m plots each) captured from a 3^2 Factorial Experiment.

	n_0	n_1	n_2	Row Total (rt)
p_0	A α n_0p_0	C β n_1p_0	B γ n_2p_0	r_0t
p_1	B β n_0p_1	A γ n_1p_1	C α n_2p_1	r_1t
p_2	C γ n_0p_2	B α n_1p_2	A β n_2p_2	r_2t
Column Total (ct)	c_0t	c_1t	c_2t	Gt

Where $r_i t = p_i(n_0 + n_1 + n_2), \quad i = 0,1,2, \quad c_j t = n_j(p_0 + p_1 + p_2), \quad j = 0,1,2,$

And gt = grand total. With respect to the latin and greek letters, they were superimposed so as to form a 3 X 3 graeco-latin square, this leads to another technique for calculating the sum of squares for the NP interaction (with 4 degrees of freedom). Since any relationship may exist between the amounts of N and P, it would be more appropriate to consider situations involving both linear and quadratic components of the regression [4] on the amount of dressing as the individual components of the main effects [5]. That is, the main effects of both N and P are, for instance, linear with the most interesting single degree of freedom from the interactions being that of N_l and P_l. From experience, since the interactions are calculated separately, one can easily see that it is this type of interaction that often approaches significance. The other three interaction components will be N_lP_q, N_qP_l and N_qP_q.

Table 3: Calculation of linear and quadratic effects for the Analysis of variance

	N_l (-1, 0, 1)	N_q (1, -2, 1)		P_l (-1, 0, 1)	P_q (1, -2, 1)
p_0	$n_2p_0-n_0p_0$	$n_0p_0-2n_1p_0+n_2p_0$	n_0	$n_0p_2-n_0p_0$	$n_0p_0-2n_0p_1+n_0p_2$
p_1	$n_2p_1-n_0p_1$	$n_0p_1-2n_1p_1+ n_2p_1$	n_1	$n_1p_2-n_1p_0$	$n_1p_0-2n_1p_1+n_1p_2$
p_2	$n_2p_2-n_0p_2$	$n_0p_2-2n_1p_2+ n_2p_2$	n_2	$n_2p_2-n_2p_0$	$n_2p_0-2n_2p_1+n_2p_2$
Sum	$N_l(p_0+p_1+p_2)$	$N_q(p_0+p_1+p_2)$	Sum	$P_l(n_0+n_1+n_2)$	$P_q(n_0+n_1+n_2)$
P_l	$N_l(p_2-p_0)$	$N_q(p_2-p_0)$	N_l	$P_l(n_2-n_0)$	$P_q(n_2-n_0)$
P_q	$N_l(p_0-2p_1+p_2)$	$N_q(p_0-2p_1+p_0)$	N_q	$P_l(n_0-2n_1+n_2)$	$P_q(n_0-2n_1+n_2)$

Note that the four components in rows 5 and 6 of columns N_l and N_q are usually the transpose of contents of columns P_l and P_q (same rows), hence all the values can be obtained from the first pair of columns (i.e. N_l and N_q).

The squares of the pertinent quantities, in table 3 with their appropriate divisors, will give an analysis of variance of the 8 degrees of freedom (d.f.) among the treatment totals into 8 single components. Recall that the entries in table 1 are individual totals over m plots, as such, the divisors are as shown in table 4 below:

Table 4: Containing the divisors for the main effects (linear) and interactions in the analysis of variance.

	N_l or P_l	N_q or P_q	$N_l P_l$	$N_l P_q$	$N_q P_l$	$N_q P_q$
Divisor	6m	18m	4m	12m	12m	36m

For instance, the divisor for N_qP_q may be worked out as follows. The three N_q figures in table 2 each have divisor, $m*(1^2 + (-2)^2 +1^2) = 6m$. Since N_qP_q is a linear function of these three figures, with coefficients 1, -2 and 1, as before, this total will have the divisor, $6m*(1^2 + (-2)^2 +1^2) = 36m$. The analysis of variance will be as shown in table 5 below:

Table 5: The analysis of variance (ANOVA) of the experiment

	d.f.	Sum (or mean) of squares
N_l	1	$(N_l(p_0+p_1+p_2))^2$ /6m
N_q	1	$(N_q(p_0+p_1+p_2))^2$ /18m
P_l	1	$(P_l(n_0+n_1+n_2))^2$ /6m
P_q	1	$(P_q(n_0+n_1+n_2))^2$ /18m
N_l P_l	1	$(N_l(p_2-p_0))^2$ / 4m
N_l P_q	1	$(N_l(p_0-2p_1+p_2))^2$ /12m
N_q P_l	1	$(P_l(n_0-2n_1+n_2))^2$ / 12m
N_q P_q	1	$(P_q(n_0-2n_1+n_2))^2$ / 36m

In table 1, the Greek and Latin letters were superimposed so as to form a (3 x 3) graeco-latin square. This square leads to another method of calculating the sum of squares for the interactions (4 d.f.). Although the method is not likely to be used for the purpose of interpretation, it has formed the basis of some devices in the construction of designs. In the square, the column totals represent the main effects of N and the row totals are those of P. In practice, the Latin letter totals are usually orthogonal to rows and columns and hence it is reasonable to suppose that they represent two of the four components of the NP interaction. Analogously, the Greek letter totals provide the remaining two components. The totals are as contained in the following table:

Table 6: Illustrating the Greek and Latin sub-totals and grand-totals:

A	B	C	Total		α	β	γ	Total
S_A	S_B	S_C	S_{ABC}		S_α	S_β	S_γ	$S_{\alpha\beta\gamma}$

Where $S_A = n_0p_0 + n_1p_1 + n_2p_2$, $S_\alpha = n_0 p_0 + n_2 p_1 + n_1 p_2$,

$S_B = n_2p_0 + n_0p_1 + n_1p_2$, $S_\beta = n_1 p_0 + n_0 p_1 + n_2 p_2$,

$S_C = n_1p_0 + n_2p_1 + n_0p_2$, $S_\gamma = n_2 p_0 + n_1 p_1 + n_0 p_2$ and ($S_{ABC} = S_A + S_B + S_C$, $S_{\alpha\beta\gamma} = S_\alpha + S_\beta + S_\gamma$).

Each quantity is by now a total of 3*12 = 36 plots. The sums of squares of deviations are obtained, individually, the usual way. For instance, the sum of squares of deviations of the Latin letter totals is $S_L = \frac{1}{36}\sum_{A,B,C}\left(S_i - \bar{S}\right)^2$. Similarly, the sum of squares of deviations of the Latin letter totals is $S'_G = \frac{1}{36}\sum_{\alpha,\beta,\gamma}\left(S_j - \bar{S}'\right)^2$, where $\bar{S}$ *and* $\bar{S}'$ are the respective, means from the Latin and Greek sub-totals. The total sums of squares of deviations $(S_L + S'_G)$ will be the same as the aggregate sum of squares for the interactions in table 5.

2.1 When the three factors have varying levels

By assuming that we have three factors; A, B, C, with each one having the respective levels; a, b, c, then the main effects will have; $(a-1)$, $(b-1)$ and $(c-1)$ degrees of freedom (or components) respectively.

2.2 Confounding in 3^n Factorial Experiments

By starting with the "least" member of the family (3^2), let us examine what confounding a 3^2 factorial will entail [6, 7].

2.2.1 Confounding of a 3^2 Factorial

By letting the symbol ij denote the treatment combination that has the i^{th} level of factor A and the j^{th} level of B (with $i, j = 1,2,3$) and keeping the main effects away from the block effects, thus confounding only the AB interaction, the size of the incomplete block will be 3 units.

Section 2.0 shows that the main effects of A and B both have two components whilst the interaction, AB has four. Besides, the main effects and the interaction can be divided into single components in many ways and the division that is most appropriate for the interpretation of the results varies from experiment to experiment. This implies that the particular component of the interaction which we should desire to confound will change from one case to the other. Confounding is more restricted in the 3^n series than the 2^n equivalent because it is based on the properties of the 3 x 3 graeco-latin square. For instance if 9 treatment combinations are set out as in the table 7 (below) where a 3 x 3 square is superimposed, then by the usual notations, one can denote the latin letters in the square by I_j, j = 1, 2, 3 respectively and the greek letters by J_i, i = 1, 2, 3 respectively.

Table 7: Using a 3 x 3 graeco-latin square to obtain an AB interaction

	b_0	b_1	b_2
a_0	(00)I_1J_1	(01)I_3J_2	(02)I_2J_3
a_1	(10)I_2J_2	(11)I_1J_3	(12)I_3J_1
a_2	(20)I3J2	(21)I_2J_1	(22)I_1J_2

Further, from table 7, it is clear that comparisons among the row totals of the square give the two components of the main effect A, while the comparisons among the column totals give the main effect of B. Let us now consider the I totals:

$I_1 = (00)+(11)+(22); \quad I_2 = (10)+(21)+(02); \quad I_3 = (20)+(01)+(12)\cdot$

With latin square, it is customary the comparisons among these totals be orthogonal to both rows and columns: that is, to the main effects of A and B. Consequently, the comparisons among the I totals must represent two of the four components of the AB interaction. An equivalent argument shows that the two remaining components of AB are obtained from the comparisons among the J totals, where

$J_1 = (00)+(21)+(12); \quad J_2 = (10)+(01)+(22); \quad J_3 = (20)+(11)+(02)\cdot$

The application of this result to confounding is shown in table 7.

Table 8: A 3 x 3 experiment with AB partially confounded

Plan (a)			Plan (b)		
Incomplete blocks			Incomplete blocks		
(i)	(ii)	(iii)	(i)	(ii)	(iii)
(00)	(10)	(20)	(00)	(10)	(20)
(11)	(21)	(01)	(21)	(01)	(11)
(22)	(02)	(12)	(12)	(22)	(02)
I_1	I_2	I_3	J_1	J_2	J_3
I components confounded			J components confounded		

In plan (a) the I components of AB are completely confounded with incomplete blocks, since the block totals have been made the same as the I totals. Here, the main effects and the J components of AB are not confounded. However, in plan (b), the J components of AB are completely confounded. The experimenter's choice for confounding between I and J is purely based on convenience since neither the I components nor its J equivalents are easy to interpret. This principle of construction can be utilized for higher factorial designs (in the 3^n series).

2.2.2 Confounding of a 3^3 Factorial

By continuing with the same symbols, such that (ijk), denotes the treatment combination $a_i b_j c_k$, there will be 27 treatment combinations altogether, the possible sizes of incomplete block will be of 3 and 9 units. Considering the plan for blocks of 9 units, only ABC need be confounded. Recall how we handled the 3^2 factorial by dividing the 9 treatments combinations into groups of three (i.e. the I and J groups), such that comparison among the group totals gave the components of AB. We now extend the principle, such that, the 27 treatment combinations for our 3^3 factorial will be divided into groups of 9, in such a way that, comparison amongst their group totals will give the components of ABC. Each set of three groups will contribute two components of ABC. Now ABC has eight components in all, hence there will be four of such sets. The AB interaction can be calculated separately for each level of C. As in the 3^2 case, we will obtain the AB interaction from it's I and J components. Table 9 (below) shows the I components for each level of C.

Table 9: The I components of AB shown for each level of C.

	C_0	C_1	C_2
I_1	$(000)+(110)+(220)$	$(001)+(111)+(221)$	$(002)+(112)+(222)$
I_2	$(100)+(210)+(020)$	$(101)+(211)+(021)$	$(102)+(212)+(022)$
I_3	$(200)+(010)+(120)$	$(201)+(011)+(121)$	$(202)+(012)+(122)$

Table 9 may also be observed as a 3 x 3 table in which each entry is the total of three treatment combinations. The row totals of the table give the I components of AB, while the column totals give the main effect of C. Further, we may take I and J totals from this table just as in table 7. These totals will be labelled $(I-I_1),(I-J_1)$ etc. because they come from the I components of AB. Hence;

$I-I_1:(000)+(110)+(220)+(101)+(211)+(021)+(202)+(012)+(122)$

$I-I_2:(100)+(210)+(020)+(201)+(011)+(121)+(002)+(112)+(222)$

$I-I_3:(200)+(010)+(120)+(001)+(111)+(221)+(102)+(212)+(022)$

Using the analogous argument to the 3^2 case, the comparisons among these totals of 9 treatment combinations must represent two of the components of the interaction of AB with C; that is, of the ABC interaction. Consequently, if we put all treatment combinations in $(I-I_1)$ into the first block, those of $(I-I_2)$ into the second and those in $(I-I_3)$ into the third, we obtain a plan which completely confounds two of the eight components of ABC and leaves all other factorial effects unconfounded. The second set of three groups of 9 treatments combinations is obtained by taking the J totals from table 9:

$I-J_1:(000)+(110)+(220)+(201)+(011)+(121)+(102)+(212)+(022)$

$I-J_2:(100)+(210)+(020)+(001)+(111)+(221)+(202)+(012)+(122)$

$I-J_3:(200)+(010)+(120)+(101)+(211)+(021)+(002)+(112)+(222)$

The remaining sets are obtained by forming a 3 x 3 table similar to table 9 for the J components of AB at each level of C.

3. Applications to the 3^n, n = 2, 3 Factorial Experiments

All other pertinent concepts will be fully illustrated in the following. Each of the two illustrations, one for each factorial experiment, is geared towards the provision of a detailed explanation on how the afore-stated theory is actually applied on the respective cases.

a. The 3^2 Factorial case

In the following, table 10, are the treatment totals for an experiment with 3 levels of Nitrogen fertilizer (N) and 3 of Phosphate fertilizer (P). The table contains data on the number of lettuce plant that emerged from the ground which are totals over m = 12 plots each. Both N and P appear to have a deleterious effect on emergence (the subscript 2 denotes the largest application). The main effects of N and P both comprise of two independent comparisons and thus have 2 degrees of freedom each. Since the amounts of N and P are somewhat in arithmetic progression, it would probably be appropriate to choose both the linear (N_l, P_l) and quadratic (N_q, P_q) components of the regression on amount of dressing as the individual components of the main effects.

Table 10: The number of lettuce plants emerging (over m=12 plots each).

	n_0	n_1	n_2	Row Total (rt)
p_0	A α 449	C β 413	B γ 326	1188
p_1	B β 409	A γ 358	C α 291	1058
p_2	C γ 341	B α 278	A β 312	931
Column Total (ct)	1199	1049	929	3177

The initial computation will appear in the following table (Table 11).

Table 11: Calculation of linear and quadratic effects for the Analysis of variance

	N_l (-1, 0, 1)	N_q (1, -2, 1)		P_l (-1, 0, 1)	P_q (1, -2, 1)
p_0	-123	-51	n_0	-108	-28
p_1	-118	-16	n_1	-135	-25
p_2	-29	+97	n_2	-14	+56
Sum	-270 (N_l)	+30 (N_q)	Sum	-257 (P_l)	+3 (P_q)
P_l	+94 (N_l)	+148 (N_q)	N_l	+94 (N_lP_l)	+84 (P_qN_q)
P_q	+84 (N_lP_q)	+78 (N_qP_q)	N_q	+148 (P_lN_q)	+78 (P_qN_q)

The left side of table 11 shows the N_l and N_q effects for each level of P, while the right side shows the P_l and P_q effects for each level of N. The column sums give the individual components of N and P main effects. For instance by considering the difference between the third and the first rows, for the N_l column, it is +94 and hence equals the linear effect of P on N (i.e. N_lP_l interaction). From the P_l column, the difference gives the linear effect of N on P_l, or the P_lN_l interaction, which is exactly the same as the N_lP_l interaction.

The other two columns provide the N_qP_l and the N_lP_q effects. The sum of the first and third rows minus twice the second row leads to the components of interaction that contain a quadratic term. It is easy to see that all four components can be obtained from either left or the right half of table 11, hence, in practice only one half is required. We have just used the computation of both halves to establish the existing symmetry amongst the components with respect to N and P. The square of these quantities, with appropriate divisors will give an analysis of variance of the 8 degrees of freedom amongst the treatment totals into 8 single components. Since an entry in the data (table 10) is the total over 12 plots, the divisors may be verified to be as shown in table 12.

Table 12: Containing the divisors for the main effects (linear) and interactions in the analysis of variance.

	N_l or P_l	N_q or P_q	$N_l P_l$	$N_l P_q$	$N_q P_l$	$N_q P_q$
Divisor	72	216	48	144	144	432

Consequently, the analysis of variance is as shown in table 13 below:

Table 13: The analysis of variance for the subdivisions of the treatment sum of squares

	Degree of freedom	sum of squares (or mean of squares)
N_l	1	1012.50
N_q	1	4.17
P_l	1	917.35
P_q	1	0.04
N_lP_l	1	184.08
N_lP_q	1	49.00
N_qP_l	1	152.11
N_qP_q	1	14.08

With an error mean of squares of about 59, the linear effects of both fertilizers are significant with no indication of curvature. The N_lP_l is significant at 10% level but not at the 5% level.

b. The case 3^3 Factorial Confounded in Blocks of 9 units

This is best illustrated with the following experiment. In this experiment, we are required to test the effects of three levels of nitrogen, three of phosphorus and three of potash on the germination of lettuce seedlings. The seeds were thoroughly mixed and divided into 108 samples, each of about 60 seeds. Each sample was planted in a copper box; 6 inches square and 1.5 inches deep containing a mixture of soil and sand. The boxes were placed in a germinator at a temperature of about 32^0C. After one week, the produce was classified as normal, abnormal, hard or dead. A summary of the numbers of normal lettuce plant is given in tables 13, 14 and 15 below. There were four replications altogether, each placed on a different shelf in the germinator. On a shelf the boxes were placed in three columns (A, B and C) of nine boxes each, each column being an incomplete block. It should be noted that the three fertilizers had a deleterious effect on emergence.

Table 14: Showing the summary of the numbers (Treatment totals) of normal lettuce plants in their respective categories.

		n_0			n_1			n_3	
	p_0	p_1	p_2	p_0	p_1	p_2	p_0	p_1	p_2
k_0	171(000)	160(010)	131(020)	174(100)	118(110)	89(120)	101(200)	102(210)	92(220)
k_1	160(001)	130(011)	99(021)	82(101)	123(111)	108(121)	120(201)	86(211)	127(221)
k_2	118(002)	119(012)	111(022)	157(102)	117(112)	81(122)	96(202)	103(212)	93(222)

Table 15: The two-way equivalent of the data in table 14 above

	n_0	n_1	n_2	Total
p_0	449	413	326	1188
p_1	409	358	291	1058
p_2	341	278	312	931
k_0	462	381	295	1138
k_1	389	313	342	1044
k_2	348	355	292	995
	1199	1049	929	3177
	p_0	p_1	p_3	Total
k_0	446	380	312	1138
k_1	371	339	334	1044
k_2	371	339	285	995
	1188	1058	931	3177

Table 16: The NPK components of the data in table 13 above

	1A	1B	1C	Total
From treatment totals	944	1102	1131	3177
From replicate 1	171	354	394	919
Difference	773	748	737	2258
	2A	2B	2C	Total
From treatment totals	1119	1113	945	3177
From replicate 2	308	251	134	693
Difference	811	862	811	2484
	3A	3B	3C	Total
From treatment totals	1025	1073	1079	3177
From replicate 3	197	232	233	662
Difference	828	841	846	2515
	4A	4B	4C	Total
From treatment totals	1104	991	1082	3177
From replicate 4	302	290	311	903
Difference	802	701	771	2274

The algorithm for the computation goes thus:

Step 1: Form the block and replicate totals and the grand total and also the totals for each treatment combinations and the three two-way tables 14, 15 and 16.

Step 2: These data enable us to calculate the total sum of squares and the sum of squares for replication, blocks within replications and for the N, P, K, NP, NK and PK factorial effects. All are obtained in the usual manner and are entered in the preliminary analysis of variance (i.e. table 18).

Step 3: There remains the calculation of the contribution from the NPK interactions. Before doing this, it may be remarked that sometimes, from the nature of the factors or from previous experience [7], there is good reason to believe that the three-factor interactions will be negligible. In this case the researcher may decide to pool the sum of squares for NPK with the error sum of squares without bordering to compute the sum of squares for NPK.. The pooled error will have 78 degree of freedom and would be obtained by subtracting the sum of squares obtained in step 2 from the total sum of squares. However, a note of caution has to be sited here because, if three-factor interactions are present the researcher may not detect them and consequently his estimate of error will be inflated. When there is doubt it is better to isolate NPK.

Step 4: In an experiment of this type, where each factor has equally spaced levels of an ingredient, it is usually advisable to examine the linear and quadratic components of the response curves. The contributions to the sum of squares of treatments are displayed in the lower section of table 17. All three fertilizers show significant linear responses, with no indication of any departure from linearity.

Table 18: Showing the analysis of variance for our 3 x 3 x 3 factorial

		Degree of freedom	sum of squares	mean of squares
Replications		3	2,041.88	
Blocks within replications		8	5,008.15	626.02
N		2	1,016.67	508.34**
P		2	917.39	458.70**
K		2	293.39	146.70
NP		4	399.27	99.82
NK		4	589.61	147.40
PK		4	212.89	53.22
NPK: confounded in replications				
1		2	25.21	12.60
2		2	64.22	32.11
3		2	6.39	3.29
4		2	198.30	99.15
Error		70	4,146.88	59.24
Total		107	14,920.25	
Subdivision of part of the treatments sum of squares				
N:	L	1	1,012.50	1,012.50**
	Q	1	4.17	4.17
P:	L	1	917.35	917.35**
	Q	1	0.04	0.04

K:	L	1	284.01	284.01**
	Q	1	9.37	9.37
NP:	L X L	1	184.08	184.08
	Rest	3	215.19	71.73
NK:	L X L	1	256.69	256.69**
	Rest	3	332.92	110.97
NPK:	L X L X L	1	59.12	59.12

Step 5: This concerns the representation of the results. Usually it will be sufficient to show the three two-way tables of means, which are derivable from the two-way tables of totals (table 15) on division by 12. Since all main effects and two factor interactions are not confounded, standard errors and t-tests for the 3 X 3 tables are obtained just as in a randomized blocks design. The principal results are that each fertilizer has produced a significant decrease in the numbers of seedlings that emerged, the decrease being adequately proportional to the amount of dressing. The significant $N_L K_L$ interaction represents the fact that the decrease in emergence from n_0 to n_2 was smaller at the k_2 level than at the k_0 level. There is an indication of a similar effect with N and P, though this is not significant.

4.0 Discussion and Conclusion

With respect to the 3^2 factorial experiments, using the data in table 10, table 6, after being fed with the appropriate quantities, gives table 19, in which the actual values of the respective totals are displayed for easy calculation of the sums of squares of the deviations for both the latin and greek letters. These add to 399.28, which is the same as the total sum of squares for the interactions in table 13, apart from rounding up differences.

Table 19: Illustrating the actual Greek and Latin sub-totals and grand-totals:

A	B	C	Total		α	β	γ	Total
1119	1013	1045	3177		1018	1134	1025	3177

Furthermore, 3^2 factorial experiments can be very complex; such is usually the case when they result from a long-term experiment. The confounding on this category of experiments further adds to its complexity, the researcher really needs to see to it that there are no mix-ups throughout the conduct of the experiment and the collation of the raw data [1]. The collation of the raw data of a 3^n factorial experiment can be very tedious [10]; to illustrate with our 3^3 factorial example, after confounding the ABC with nine units per block, the raw data actually looked like the contents of tables 19 through 22, it was through summarizing that tables 14 through 16 were obtained.

Table 20: Showing the data entries (number of lettuce plants emerging), in our 3^3 factorial example (rep. I)

REPLICATE I

Blocks		A		B		C	Total
	npk	No.	npk	No.	npk	No.	
	012	11	201	42	111	46	
	122	11	121	20	102	48	
	220	13	210	24	221	58	
	202	12	011	38	001	53	
	101	11	112	39	010	54	
	021	30	100	61	212	37	
	000	41	002	40	022	41	
	110	21	222	46	120	25	
	211	21	020	44	200	32	
Total		171		354		394	919

Table 21: Showing the data entries (number of lettuce plants emerging), in our 3^3 factorial example (rep. II)

REPLICATE II

Blocks		A		B		C	Total
	npk	No.	npk	No.	npk	No.	
	121	39	100	32	010	26	
	220	34	012	37	120	16	
	102	31	210	37	200	06	
	201	34	221	33	021	13	
	000	40	202	27	211	12	
	022	31	020	21	222	12	
	212	26	111	12	002	16	
	011	35	001	30	112	19	
	110	38	122	22	101	14	
Total		308		251		134	693

Table 22: Showing the data entries (number of lettuce plants emerging), in our 3^3 factorial example (rep. III)

REPLICATE III

Blocks		A		B		C	Total
	npk	No.	npk	No.	npk	No.	
	101	26	020	19	220	19	
	210	17	211	25	122	19	
	221	21	121	18	100	42	
	000	38	110	27	201	11	
	112	22	222	20	010	36	
	011	19	001	60	111	37	
	120	18	012	30	002	29	
	202	22	102	28	021	27	
	022	14	200	25	212	13	
Total		197		232		233	662

Table 23: Showing the data entries (number of lettuce plants emerging), in our 3^3 factorial example (rep. IV)

REPLICATE IV

Blocks		A		B		C	Total
	npk	No.	npk	No.	npk	No.	
	112	37	020	47	222	15	
	121	31	002	33	201	42	
	010	44	221	15	012	41	
	022	25	122	29	210	24	
	100	39	011	38	000	54	
	211	28	212	27	111	28	
	202	35	200	38	120	30	
	001	37	101	31	102	50	
	220	26	110	32	021	29	
Total		302		290		311	903
					Grand total	3177	

Consider the two components of NPK that are confounded in replication I. The contribution of these components to the sum of squares for NPK must be calculated from the remaining 3 replicates, in which they are not confounded with blocks. The totals needed are shown in table 15. From the treatment totals, compute the total (i.e. 944) of the 9 treatment combinations that appear in block 1A. That is;

$$Block \quad 1A = 012+122+220+202+101+021+000+110+211$$
$$=119+81+92+96+82+99+171+118+86=944$$

In the same way 1102 and 1131 are obtained for the totals over all treatments that appear in blocks 1B and 1C respectively. Under these three figures (i.e. 944, 1102 and 1131) are placed the respective totals for blocks 1A, 1B and 1C. Through subtraction we obtain the totals 773, 748 and 737 (the totals of the groups of treatment combinations taken over replications 2, 3 and 4). The sum of squares of deviations of these quantities from their mean is divided by 27, since each figure contains 27 observations. The result 25.21 is the contribution of the two components of NPK to the sum of square for NPK. The remaining six components are found similarly from replications 2, 3 and 4. All eight components could be computed in one step as:

$$\frac{1}{27}\left\{(773)^2+(748)^2+...+(701)^2+(771)^2\right\}-\frac{1}{81}\left\{(2258)^2+...+(2274)^2\right\}$$

The Error sum of squares, with 70 degrees of freedom is obtained through subtraction.

With respect to the two-factor interactions, the "linear by linear" components have been isolated for NP and NK, section 3.1 contain the details for the method. In the case of PK, it was not worthwhile because the total sum of squares for PK (i.e. 212.89) is not large enough to allow any single component to be significant. The linear by linear component is significant for NK but not for NP.

Whenever the linear by linear components of the two-factor interactions are large, it usually becomes desirable that the linear by linear by linear (L X L X L) for the three-factor interaction be isolated from it.

Finally, the individual 27 treatment totals or means cannot be used as they are for interpretative purposes because they each contain some quantity of block effect. A table of these totals or means will not be necessary except if the three-factor is under study. However, to obtain such a table, we could adjust each total so as to remove the inherent block effects. Each block effect is first estimated. Here, we do not use the observed block mean because that also contains treatment effects. Hence, the least square estimate of any block effect is;

$$\frac{1}{27}\left\{4(block \quad total)-(total \quad of \quad treatments \quad appearing \quad in \quad the \quad block)\right\}$$

The required data for this computation is contained in Table 15. Thus, for block 1A, the estimated effect is $\frac{1}{27}\left[4(171)-(944)\right]=-9.6$.

The block effects are given in Table 24.

Table 24: Showing all the block effects

	Block		
Replication	A	B	C
1	-9.6	11.6	16.5
2	4.2	-4.0	-15.1
3	-8.8	-5.4	-5.4
4	3.9	6.3	6.0

In order to adjust any treatment total, we note the four blocks in which it features and compute the sum of the effects for these four blocks. This quantity is subtracted from the unadjusted total to give the adjusted total. For the purpose of illustration, consider $n_1p_1k_0$, which appears in blocks 1A, 2A, 3B and 4B, the adjusted total for it is; $118-\{-9.6+4.2-5.4+6.3\}=122.5$. In order to obtain the adjusted mean we usually divide this quantity by 4 (i.e. $\frac{122.5}{4}=30.63$).

References

[1] Box, G E P, Hunter, J S and Hunter W G (2005). Statistics for Experimenters (Design, Innovation and Discovery) Second Edition. A John Wiley and Sons, Inc., Publication Hoboken, New Jersey.

[2] Cochran, W G and Cox, G M (1957) Experimental Designs (second edition). A Wiley publication in Applied Statistics. John Wiley and Sons, Inc, London

[3] Hinkelmann, K and Kempthorne, O (2007). Design and Analysis of Experiments (second edition) Vol. 1. A John Wiley and Sons, Inc., Publication. Hoboken, New Jersey.

[4] Wong, W K (1994). Comparing Robust Properties of A, D, E and G-optimal Designs. Computational Statistics and Data Analysis 18, 441-448, North-Holland.

[5] Vicente, G, Coteron, A, Martinez, M and Aracil, J (1998). Application of the Factorial Design of Experiments and Response Surface Methodology to optimize Biodiesel Production. Industrial Crops and Products 8, 29-35.

[6] Fontana, R and Sampo, S (2013). Minimum-Sized Mixed Level Orthogonal Fractional Factorial Designs Generation: A SAS-Based Algorithm. Journal of Statistical Software Vol. 53, Issue 10 (http://www.jstatsoft.org/).

[7] Tan, Y and Zhang, R (2013). Construction of Two-level Blocked Minimum Lower Order Confounding Designs. Pak. J. Statist. Vol. 29(3), 351-367.

[8] Phoa, F K H (2009). Analysis and Construction of Nonregular Fractional Factorial Designs. A Dissertation in the Department of Statistics, University of California, Los Angeles.

[9] Aoki, S (2010). Some Optimal Criteria of Model-Robustness for Two-level Non-regular Fractional Factorial Designs. Ann. Inst. Stat. Math. 62: 699-716 (Doi 0.1007/s10463-010-0292-7).

[10] Jaynes, J, Ding, X, Xu, H, Wong, W K and Ho, C(2013). An Application of Fractional Factorial Designs to study Drug Combinations. Stat. Med. January 30; 32(2) (Doi:10.1002/sim.5526).