19 March 1945

Dear Mr Lyle,

I have just been looking at your paper on the design and analysis of experiments and wondered whether it would not be useful to your men to do the kind of thing which I have illustrated in the <u>Design of Experiments</u>.

In your first example, one can put down the data shown in Table 1, which is complete save for the 18 degrees of freedom for e ror within cells, for which the corresponding sum of squares is, from Table 13, .06260, giving a mean square .003[illegible] Table 1 shows, however, only the means of the triplicate trials, and I sho ld myself perform the following process with the totals, which would give sums of squares and mean squares nine times those which I shall get, i.e. they will need to be divided by 3, instead of requiring to be multiplied by 3 in order to be comparable with the mean square for error within cells.

Calling the three lines of the t ble <u>a</u>, <u>b</u> and <u>c</u>, I make three new lines marked $a+b+c$, $a-c$, and $a-2b+c$ respectively. The sums of the squares of the coefficients of a, b and c in these expressions are 3, 2 and 6, which are noted in the three lines. Next one repeats the same process horizontally, as though the columns are now designated a, b and c, so obtaining nine quantities, of which one 2.21, is irrelevant to the analysis, being merely the grand total of the nine entries. The other 8 correspond with the 8 degrees of freedom into which variation between cells may be analysed. The entry under Sum of Squares for each of these degrees of freedom is in fact the square of the corresponding number just obtained, divided by an appropriate divisor obtained from the numbers 3, 2, 6 noted above, and being in fact simply

(9)	6	18
6	4	12
18	12	36

which are the divisors corresponding with the nine numbers obtained earlier, vis.

(2.21)	1.13	0.39
0.71	0.27	0.05
0.03	0.23	0.09.

Squaring these and dividing, one has

	$.21281\dot{8}$	.00845
$.08401\dot{6}$	.018225	$.000208\dot{3}$
.00005	$.004408\dot{3}$	.000225

Five of these numbers, as you will see, appear duly multiplied by 3 in your Table 13. The other three are pooled in your entry for error between cells, but there is a certain advantage in looking at them individually before deciding they are practically pure error. Indeed the interaction between the linear effect of temperature and the quadratic effect of concentration is high enough, relative to the mean square with in cells, to arouse a certain suspicion of reality. The other two are obviously trifling.

Of course, equivalent results are the results of equivalent processes, but the operations on the enclosed sheet are often found quite enlightening in similar work.

For a 5 x 5 design, I should be inclined to pick out the same eight linear and quadratic effects of each factor, and their interactions, and to derive the remaining 16 degrees of freedom between cells by subtraction. Here and for larger numbers one can use coefficients from the Table of orthogonal polynomials which Yates and I have published in Statistical Tables, i.e. for the five the linear term 2a + b − d − 2e with divisor 10, and the quadratic term 2a − b −2c −d + 2e with divisor 14. Of course, however, in this case, if the 16 residual degrees of freedom seem to have more than error in them, one could explore further with the cubic and even with the biquadratic term given in the same Table.

Yours sincerely,

		a	0.78	0.42	0.23
		b	0.52	0.32	0.22
		c	0.40	0.20	0.12
Total	3	a+b+c	1.70	0.94	0.59
Linear	2	a-c	0.38	0.22	0.11
Quadratic	6	a-2b+c	0.14	-0.02	-0.09

Total	Linear	Quadratic
3	2	6
a+b+c	a-c	a-2b+c
2.21	1.13	0.39
0.71	0.27	0.05
0.03	0.23	0.09