SAMPLING, BLOCKING, AND MODEL CONSIDERATIONS FOR THE r-ROW BY c-COLUMN EXPERIMENT DESIGNS

BU-550-M | by | Revised July, 1975

W. T. Federer

Use "Summary" on p. 17 for "Abstract".

Paper No. BU-550-M in the Biometrics Unit Mimeo Series, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14853.

Cornell University
Ithaca, New York

Sampling, Blocking, and Model Considerations for the r-row by c-column Experiment Designs

W. T. Federer

1. Introduction

In a previous paper (Federer (1975)), the sampling and population structures, considerations involved in blocking, form of treatment response, and some practical considerations were presented for unblocked and blocked experiment designs. These same topics will now be considered for r-row by c-column experiment designs including the Latin square, Youden, and other designs for two-way elimination of heterogeneity. The same definitions and symbolism presented in the previous paper will be used herein, and hence will not be repeated.

We first consider the population and subpopulation distributional and sampling structures for an r-row by c-column experiment design prior to application of the v treatments. Then, the additive and multiplicative form of treatment responses are considered. An alternative sampling structure is discussed; this structure simulates many experiments in practice but the resulting statistical inferences may be in question.

In the third section, several forms of r-row by c-column designs are presented; examples of the Latin square, the Youden, F-squares, simple change-over, and several other types of experiment designs are given to demonstrate the wide range of such designs. A wide variety of uses to which these designs have been put is presented in the fourth section.

Statistical texts are prone to presenting a single statistical analysis for all Latin square and Youden designs. Depending upon the form of treatment response and model, many statistical analyses should be considered for experiments as diverse as those presented in section four. It is highly unlikely that all experiments can be analyzed in the same cookbook manner. It is essential to obtain the appropriate statistical analysis for each experiment if the statistical and subject matter inferences are to have any validity.

2. The sampling structure for row-column experiment designs and forms of response

Suppose that a population of experimental units exists such that there are RC subpopulations of the following nature:

"rows" = 1st source of variation	"columns" = 2nd source of variation						
	1	2	3	...	i	...	C
1							
2							
3							
⋮							
h							
⋮							
R							

R and C could be finite or infinite. The hi^{th} subpopulation consists of a large number (infinite) of experimental units. Denote the arithmetic average over all subpopulations of experimental units as $\mu_{..}$, the arithmetic average of the i^{th} "column" over all "rows" as $\mu_{.i}$, and the arithmetic average of the h^{th} "row" over all "columns" as $\mu_{h.}$. Then, for additive effects and non-interactive "rows" and "columns", the arithmetic average of the h^{th} "row" and i^{th} "column" is

$$\mu_{h.} + \mu_{.i.} - \mu_{..} = \mu_{..} + (\mu_{h.} - \mu_{..}) + (\mu_{.i} - \mu) = \mu_{..} + \rho_h + \gamma_i \ . \tag{2.1}$$

A randomly selected observation from the hi^{th} subpopulation could be described as having the following response and parametric structure:

$$Y_{hi} = \mu + \rho_h + \gamma_i + \epsilon_{hi} \,, \tag{2.2}$$

where $\epsilon_{hi} = Y_{hi} - \mu - \rho_h - \gamma_i$ and $\mu_{..} = \mu$.

Further, suppose that the randomly distributed ϵ_{hi} are independently and identically distributed with zero mean and common variance σ^2_ϵ. Then, <u>IF</u> a treatment, say j, effect is denoted as τ_j and is additive irrespective of which "row" and "column" has been selected, then the model equation for the observation, yield, or response obtained from applying treatment j to a randomly selected experimental unit from subpopulation hi is

$$Y_{hi} + \tau_j = Y_{hij} = \mu + \tau_j + \rho_h + \gamma_i + \epsilon_{hij} = \mu_{..j} + \rho_h + \gamma_i + \epsilon_{hij} \tag{2.3}$$

where the subscript on ϵ_{hij} denotes that treatment j has been applied to a randomly selected experimental unit from subpopulation hi. Note that the ϵ_{hij} have zero mean and variance σ^2_ϵ. The addition of a treatment has done nothing but raise or lower the value of Y_{hi} by τ_j. $\mu_{..j}$ is the average value resulting from applying treatment j to every individual in every one of the RC subpopulations.

One of many possible alternate models would be to have the treatment effects affect the experimental units multiplicatively as follows:

$$Y^*_{hij} = \tau^*_i(\mu + \rho_h + \gamma_i + \epsilon_{ij}) = \tau^*_j(\mu_{h.} + \mu_{.i} - \mu + \epsilon_{hi}) = \mu^*_{h.(j)} + \mu^*_{.i(j)} - \mu^*_{..j} + \epsilon^*_{hij} \tag{2.4}$$

where $\mu^*_{h.(j)}$ is the arithmetic average obtained by applying treatment j to every individual of all subpopulations in "row" h, $\mu^*_{.i(j)}$ is the mean of every individual receiving treatment j in all subpopulations in "column" i, and $\mu^*_{..j}$ is the mean obtained from applying treatment j to every experimental unit in the RC subpopulations. Then, $\tau_j^* = \mu_{..j}^*/\mu$. This is the situation presumed when treatment means are reported in percent of a standard or control treatment. Note also that the ϵ_{hij}^* have zero mean and variance $\sigma^2_{\epsilon^* j} = (\tau_j^*)^2\sigma^2_\epsilon$.

The terms "rows" and "columns" are generic symbols denoting two sources of variation and need not refer to rows and columns of a lattice or of a rectangle. Hereafter we shall drop the quotes but not the general meaning. To obtain an r-row by c-column design of rc experimental units (e.u.'s) for an upcoming experiment, we should

(i) select the v treatments for the experiment,

(ii) obtain a simple random sample of r rows,

(iii) obtain a simple random sample of c columns,

(iv) randomly select one experimental unit (or set of k units) from each selected subpopulation, and

(v) randomly assign the v treatments to the rc experimental units using the constraints of the particular experimental design selected.

An alternate sampling structure and distribution is to conceive of units of size r e.u.'s by c e.u.'s, and to consider these units of rc e.u.'s as coming from a single population composed of such groupings into units. This population has a mean μ and a constant variance σ^2 among such units of size rc e.u.'s. The experimenter randomly selects <u>one</u> of these units, say the g^{th} one, and lays out an r-row by c-column design on this unit. The variational model is then considered to be

$$Y_{ghi} = \mu_g + \rho_{gh} + \gamma_{gi} + \epsilon_{ghi}, \tag{2.5}$$

prior to the addition of treatment j; let the ϵ_{ghi} be independently and identically distributed with mean zero and variance σ^2_ϵ. Suppose that equation (2.3) and model is appropriate after applying the j^{th} treatment to the ghi^{th} unit from the population. Note that even though only a single element has been obtained from the population, differences between estimated treatment effects are validly estimated because the difference is independent of the g^{th} unit.

The above sampling situation is more nearly akin to what actually happens in experimentation. A geographically connected or spatially connected set of rc e.u.'s is obtained and an experiment is conducted. For example, a piece of land is selected and the rows and columns are set up to account for suspected or known gradients in the experimental area; the area is divided into r rows and c columns and conforms to the description above rather than the one in the first part of this section.

It should be noted that if this latter situation prevails, serious consideration should be given to the analyses described by Kempthorne (1952) as randomization tests. Also, the variance more than likely will depend upon the units of rc e.u.'s and will be $\sigma^2_{\epsilon g}$ rather than σ^2_{ϵ}. The effect on interval estimation should be noted here. In addition, if model (2.4) holds, the experimenter is faced with making inferences from a sample of size one. All such items as the above should be considered prior to conducting an experiment to ascertain how it should be designed in order to make valid inferences.

3. Some row-column designs

We shall discuss only experimental designs in which the row, column, and treatment effects are orthogonal and which have responses of the form of equation (2.3). This means that arithmetic means are used to summarize the data when the linear, additive model holds. One of the simplest and most used r-row by c-column designs for v treatments replicated b times is the Latin square design. Here $r = c = v = b = n$ is the order of the Latin square and n^2 observations are obtained. Each of the n treatments occurs <u>once</u> in each row and <u>once</u> in each column. For $n = 2, 3, 4$, and 5, the following are systemmatically arranged Latin squares:

n = 2

A	B
B	A

n = 3

A	B	C
C	A	B
B	C	A

n = 4

A	B	C	D
D	A	B	C
C	D	A	B
B	C	D	A

n = 5

A	B	C	D	E
E	A	B	C	D
D	E	A	B	C
C	D	E	A	B
B	C	D	E	A

B	A
A	B

A	B	C
B	C	A
C	A	B

A	B	C	D
B	A	C	D
C	D	A	B
D	C	B	A

A	B	C	D	E
B	A	D	E	C
E	C	A	B	D
D	E	B	C	A
C	D	E	A	B

The treatments are denoted by Latin letters, and the array of letters is square. Hence, the term Latin square. If one had used Greek letters, we could have used the term Greek square instead.

One possible randomization procedure frequently used is described on page 117 of Federer (1973) and page 207 of Cox (1958). (See also Federer (1955) and Kempthorne (1952).) Another randomization procedure useful for any r-row by c-column design is:

(i) randomly allot the treatments to the v letters,

(ii) randomly allot the v treatment letters to the experimental units in the c columns in row 1 of the experiment,

(iii) randomly allot the v treatments to columns in row 2 of the experiment, except to make certain that the number of times any treatment letter occurs in a column does not exceed the number of times allowed (once for a Latin square),

(iv) randomly allot the v treatments except for the proviso in (ii), and

(v) continue the process until all rows are allotted.

This procedure allows any possible plan to be selected equally frequent.

Consider another class of row-column designs for which there are v rows, v treatments, and vs = c columns. These are the simple change-over designs. Examples for v = 2, c = 6, and v = 3, c = 6 are:

v = 2

row	column 1	2	3	4	5	6
1	A	A	B	A	B	B
2	B	B	A	B	A	A

v = 3

row	column 1	2	3	4	5	6
1	A	A	B	C	B	C
2	B	C	A	B	C	A
3	C	B	C	A	A	B

An example of another class of designs is:

row	Blocks (columns) 1			2			3		
1	A	B	C	G	H	I	D	E	F
2	D	E	F	A	B	C	G	H	I
3	G	H	I	D	E	F	A	B	C

Note that all treatments occur once in each row and all treatments occur once in each column or block; the row-column intersection contains more than one experimental unit.

Another class of orthogonal row-column designs is one in which the proportion of times any treatment occurs in a row or a column is constant. Thus, treatment j occurs λ_{rj} times in each row and λ_{cj} times in each column; if r = c, then $\lambda_{cj} = \lambda_{rj}$. If $\lambda_{rj} = \lambda_{cj} = \lambda$ for all j, then designs similar to the simple-change-over result. If $\lambda = 1$, then the Latin square results. Designs with differing $\lambda_{rj} = \lambda_{cj} = \lambda_j$ are called F-squares. Some examples for r = c = 4 are:

$\lambda_A=2,\ \lambda_B=1,\ \lambda_C=1$

A	B	C	A
A	A	B	C
C	A	A	B
B	C	A	A

$F(A^2,B,C)$

$\lambda_A=2=\lambda_B$

A	B	A	B
B	A	B	A
A	B	A	B
B	A	B	A

$F(A^2,B^2)$

Examples for r = c = 5 are:

$\lambda_A=3,\ \lambda_B=\lambda_C=1$

A	B	C	A	A
A	A	B	C	A
A	A	A	B	C
C	A	A	A	B
B	C	A	A	A

$F(A^3,B,C)$

$\lambda_A=2=\lambda_B,\ \lambda_C=1$

A	A	B	B	C
C	A	A	B	B
B	C	A	A	B
B	B	C	A	A
A	B	B	C	A

$F(A^2,B^2,C)$

$\lambda_A=4,\ \lambda_B=1$

A	A	A	A	B
B	A	A	A	A
A	B	A	A	A
A	A	B	A	A
A	A	A	B	A

$F(A^4,B)$

Many nonorthogonal n-row by n-column designs for v treatments with not too difficult statistical analyses are available. For example, consider the following two:

v = 9 treatments

row	column 1	2	3	4	5	6
1	A	B	D	E	G	H
2	B	C	E	F	H	I
3	C	A	F	D	I	G
4	G	H	A	B	D	E
5	H	I	B	C	E	F
6	I	G	C	A	F	D

v = 4 treatments

row	column 1	2	3	4	5	6
1	A	B	C	D	A	C
2	D	A	B	C	B	D
3	C	D	A	B	D	A
4	B	C	D	A	B	C
5	A	B	D	C	A	B
6	B	C	A	D	C	D

Thus, we may put 9 treatments replicated 4 times each in a 6 x 6 square, or we may put 4 treatments replicated 9 times in a 6 × 6 square. Another design for 9 treatments in a 6-row by 9-column rectangle would be:

row	column 1	2	3	4	5	6	7	8	9
1	A	B	C	D	E	F	G	H	I
2	B	C	A	E	F	D	H	I	G
3	C	A	B	F	D	E	I	G	H
4	D	E	F	G	H	I	A	B	C
5	E	F	D	H	I	G	B	C	A
6	F	D	E	I	G	H	C	A	B

Many other variations are possible. One such variation with a relatively simple statistical analysis is the so-called Youden design (sometimes called a Youden square even though it is a rectangular array). A Youden design may be made by adding or by deleting a row to an ordinary Latin square. Other Youden designs are possible. For example, Youden designs for v = 7 and 13 treatments are:

v = 7 treatments

row	column 1	2	3	4	5	6	7
1	A	B	C	D	E	F	G
2	B	C	D	E	F	G	A
3	D	E	F	G	A	B	C

v = 7 treatments

row	column 1	2	3	4	5	6	7
1	C	D	E	F	G	A	B
2	E	F	G	A	B	C	D
3	G	F	A	B	C	D	E
4	G	A	B	C	D	E	F

v = 13 treatments

row	column 1	2	3	4	5	6	7	8	9	10	11	12	13
1	A	B	C	D	E	F	G	H	I	J	K	L	M
2	C	D	E	F	G	H	I	J	K	L	M	A	B
3	I	J	K	L	M	A	B	C	D	E	F	G	H
4	M	A	B	C	D	E	F	G	H	I	J	K	L

Youden designs have the property that all treatments occur in each row and that every pair of treatments occurs together in the column an equal number, λ, of times. $\lambda = 1$ for the first and third designs above, and $\lambda = 2$ for the 4-row by 7-column design above.

Many other designs with two-way elimination of heterogeneity are possible. Some of these are listed in sections XIII-3, XIV-1, XIV-2, and XV-2 of Federer (1955). Some additional designs are presented by Federer and Raghavarao (1975) and Federer *et al.* (1975). These designs are for use in varietal and drug screening trials.

4. Some practical applications of two-way blocked designs

An experiment (Bliss and Rose (1940)) designed as a 4 x 4 Latin square, was one in which the treatments were 4 preparations of a dosage of an extract of parathyroid, the rows were different dogs, and the columns were different days. The treatments were labeled U_1, U_2, S_1, and S_2. The dosages were given to 4 dogs at 4 different times as follows:

	Day (column)			
Dog (row)	1	2	3	4
1	S_1	S_2	U_2	U_1
2	U_2	U_1	S_1	S_2
3	S_2	S_1	U_1	U_2
4	U_1	U_2	S_2	S_1

The outcome or measurement was in terms of mg.-percent serum calcium in the blood.

As a second example (Thomson (1941)), 4 groups of children (columns), 4 different word lists (rows), and 4 methods of testing spelling (treatments): multiple choice (MC), second dictation (SD), wrongly spelled words (WS), and skeleton words (SK), were designed in a Latin square as follows:

Word List (row)	Groups of Children (column) I	II	III	IV
1	MC	SD	WS	SK
2	SK	MC	SD	WS
3	WS	SK	MC	SD
4	SD	WS	SK	MC

with the response being percentage of correctly spelled words.

As a third example (Maxwell (1958)), three forms (AA, AB, and B) of the Nufferno speed test with three students and three times of day in a Latin square arrangement were used as follows:

Time of day	Student (female, 20 years old) 1		2		3	
morning	AA	0.76	AB	0.86	B	1.12
afternoon	AB	0.82	B	1.16	AA	0.74
evening	B	0.98	AA	0.83	AB	0.83

The responses (test scores) are given above for this experiment. The mean responses for the three treatments were: AA - 0.78, AB - 0.84, and B - 1.09.

Another use for the Latin square design is in bridge games or other tournaments where each couple plays every other couple (round robin tournaments). For three evenings of bridge, three games per evening, and four couples (two tables), the following illustrates the design:

Game number row	Evening 1	2	3
1st game	12 34	13 24	14 23
2nd game	14 23	12 34	13 24
3rd game	24 13	14 23	12 34

For example, for the first game on evening 1, couple 1 plays couple 2 and couple 3 plays couple 4; then, for the second game on evening 1, couple 1 plays couple 4 and couple 2 plays couple 3; and for the third game on evening 1, couple 1 plays couple 3 and couple 2 plays couple 4. Each couple plays every other couple on each evening. Likewise, each couple plays every other couple in the first game of an evening, etc. The couples are randomly numbered and the particular rows in the above Latin square are randomly assigned to the order of games. (Also, see Cochran (1971)).

The results of an agricultural trial described in the 1932 Report of Rothamsted Experimental Station comprises still another example of a Latin square design. The six treatments represent different quantities of nitrogeneous and phosphatic fertilizers applied to young potato plants. The response or measurement is in pounds of potatoes harvested from a plot of ground. A rectangular area of land was selected. Since soil gradients in two directions were suspected, the following layout in a 6 x 6 Latin square design was utilized:

row	column 1	2	3	4	5	6
1	E-633	B-527	F-652	A-390	C-504	D-416
2	B-489	C-475	D-415	E-488	F-571	A-282
3	A-384	E-481	C-483	B-422	D-334	F-646
4	F-620	D-448	E-505	C-439	A-323	B-384
5	D-452	A-432	B-411	F-617	E-594	C-466
6	C-500	F-505	A-259	D-366	B-436	E-420

(The data and statistical analysis for the above experiment are discussed in Fisher (1966), Tables 9, 30, 31, and 32.)

Some examples of other types of situations in which Latin square or Latin rectangle designs have been used are:

row	column	treatment	response
day of week	operator inserting primer in explosives	team mixing an explosive	proportion of defective shells
turnip plant	size of leaf	storage time of leaf	moisture content in percent
grocery store	day of week	method of packaging apples	sales in pounds per 100 customers
hour of day	day of week	light intensity	bioelectric potential difference between two points on a plant
segment of a planted row	order within segment	pruning method	total yield of grapes in pounds
row of a rectangular area of land	column of a rectangular area of land	cabbage variety	total pounds of marketable heads per area
cow	6-week period	nutritional diet	pounds of butterfat per 6-week period
rabbit	date of injection	level of insulin	blood sugar content

(continued)

row	column	treatment	response
traffic level and load	order of treatment on road	material for a hard surfaced road	deterioration of road
automobile	wheel location	brand of tire	tread measurement
target	order of bomber group over target	bomber group	proportion of bombs on target
day of week	strain of pigeon	drug and day length	amount of sexual activity
thermometer	battery cell	day of measurement	temperature
time period	order of measurement on electroplated panel	unknown or a standard	radioactivity relative to a standard
plant	position of leaf	virus solution	number of lesions
pen	size of pig litter	diet supplement	weight at 54 days of age
set of numbers	order of calculation	calculating machine	time in seconds to compute a sum of squares
area of wheat of 80 plants	order of estimate	an inspector	mean height of 8 "representative" shoots of wheat
time of day plus observer	day of week	poisonous injection in femoral vein of cat	length of time to death
machine	lot of material	method of weaving cotton cloth	length of wear of cloth
method of curing tea leaves	date of harvest	manurial treatment	score by a panel of tea tasters
color of light	order of presentation to subjects	level of illumination	number of eye blinks per minute

The above represents but a small sample of uses of Latin square and Latin rectangle designs, but these should be sufficient to indicate the diversity of usage. In all situations the experimenter should become familiar with the sampling structure of his population prior to conducting an experiment and applying a treatment; the nature of the response when treatments are applied must also be known in order to make correct statistical summarizations and statistical inferences. (See, e.g., Cox (1958), Federer (1955, 1975), Fisher (1966), and many other texts for examples of experiments in a row-column design.)

5. Statistical analyses

From the diverse list of applications presented in the previous section, it is illogical to believe that one model, namely (2.3), would hold for all of them. However, this is what was assumed for the analyses of the data, and of course, this was the only analysis most experimenters had ever been taught in a statistics course. One exception to the following standard analysis of variance table on either transformed or untransformed data for a Latin square design:

Source of variation	Degrees of freedom
Total	n^2
Correction for the mean	1
Among rows	n-1
Among columns	n-1
Among treatments	n-1
Remainder which is equated to error	(n-1)(n-2)

is a one degree-of-freedom sum of squares for nonadditivity as developed by Tukey (1955). (See also section 11.20 in Snedecor and Cochran (1967) for a discussion and application of the procedure.) For this situation it is postulated that

$$E[Y_{hij}] = \mu + \rho_h + \gamma_i + \tau_j + c_1(\rho_h\gamma_i + \rho_h\tau_j + \gamma_i\tau_j)$$

and that the alternate hypothesis for additivity is

$$H_A: c_1 \neq 0 , \tag{5.1}$$

where the symbols are as defined in equation (2.3). Note that one could have added the term $c_2\rho_h\gamma_i\tau_j$ to the above expectation, but c_2 is assumed to be zero for the test. Alternatively, one could consider four different hypotheses for the four different types of nonadditivity involved. Note that

$$(5.2) \quad \mu\rho_h\gamma_i\tau_j = \mu_{h..}\mu_{.i.}\mu_{..j}/\mu^2 = \mu + \rho_h + \gamma_i + \tau_j + (\rho_h\gamma_i + \rho_h\tau_j + \gamma_i\tau_j)/\mu + \rho_h\gamma_i\tau_j/\mu^2 .$$

Nair (1975) has discussed this problem and has given separate tests for each type of nonadditivity.

Another nonstandard statistical analysis for data from an experiment designed as a Latin square is the one given by Cox (1956) wherein different gradients are considered for either the rows or the columns. A generalization of this would consider differential gradients in both rows and columns. Still another statistical analysis would involve the use of a covariate rather than a blocking variable to eliminate the heterogeneity present for that source of variation; this could be used when the value of the covariate was present such as was, for example, for the data given in Example VI-1 of Federer (1955). All of these analyses for Latin square designs could be extended to the general r-row by c-column designs. Statistical analyses under the model given in equation (2.4) have been presented by Nair (1975). These are but a few of the possible models and statistical analyses for r-row by c-column designs. Two problems that require resolution are the determination of the correct error structure and the estimation of estimable functions under models other than (2.3). It appears that considerable work is required to resolve all such problems associated with the statistical analysis of these designs.

In much the same manner as for blocked designs (Federer (1975)), a test of model (2.3) with model (2.4) as the alternative hypothesis, could be obtained by computing the treatment means adjusted for any nonorthogonality from rows and columns and the sums of squares $\sum_{h,i} (\hat{\epsilon}_{hij})^2$ of residuals for each treatment j, by ranking the means and sums of squares, and by computing Spearman's rank order correlation.

6. Summary

Two different sampling structures, prior to the addition of treatments, for experiments designed as r-row by c-column designs are discussed. Their relations to actual applications are considered together with additive and multiplicative treatment responses. Many types of experiments in which the Latin square design has been used are presented. Types of available statistical analyses are briefly discussed.

References

BLISS, C. I. and C. L. ROSE, 1940: The assay of parathyroid extract from the serum calcium of dogs, American Journal of Hygiene 31, 79-98.

COCHRAN, R. S., 1971: Designs with redeeming social aspects for evenings of social bridge, The American Statistician 25(2), 12-15.

COX, C. P., 1956: The analysis of Latin square designs with individual curvatures in one direction, Journal of the Royal Statistical Society, Series B, 20, 193-204.

COX, D. R., 1958: Planning of Experiments. John Wiley and Sons, Inc., New York, London, and Sydney.

FEDERER, W. T., 1955: Experimental Design - Theory and Application. The Macmillan Company, New York. (Indian Edition by Oxford & IBH Publishing Co., 1967.)

FEDERER, W. T., 1973: Statistics and Society. Marcel Dekker, Inc., New York.

FEDERER, W. T., 1975: Sampling, blocking, and model considerations for the completely randomized, randomized complete block, and incomplete block experiment designs, BU-548-M in the Mimeo Series of the Biometrics Unit, Cornell University.

FEDERER, W. T., R. C. NAIR, and D. RAGHAVARAO, 1975: Some augmented row-column designs, Biometrics 31(2), 361-373.

FEDERER, W. T. and D. RAGHAVARAO, 1975: On augmented designs, Biometrics 31, 29-35.

FISHER, R. A., 1966: The Design of Experiments, 8th edition (1st edition 1935), Oliver and Boyd, Ltd., Edinburgh.

KEMPTHORNE, O., 1952: Design and Analysis of Experiments. John Wiley and Sons, Inc., New York.

MAXWELL, A. E., 1958: Experimental Design in Psychology and the Medical Sciences. Methuen and Co., Ltd., London; John Wiley and Sons, Inc., New York.

NAIR, R. C., 1975: Some contributions to the theory of nonadditivity in data analysis, Ph.D. Thesis, Cornell University.

SNEDECOR, G. W. and W. G. COCHRAN, 1967. Statistical Methods, 6th edition. The Iowa State University Press, Ames, Iowa, U.S.A.

THOMSON, G. H., 1941: The use of the Latin square in designing educational experiments, British Journal of Educational Psychology 11, 135-137. (Data given by S. D. Nisbet (1939): Non-dictated spelling tests, British Journal of Educational Psychology 9, 20-44.)

TUKEY, J. W., 1955: Query 113, Biometrics 11, 111-113.