Latin squares: Some history, with an emphasis on their use in designed experiments

R. A. Bailey
University of St Andrews
QMUL (emerita)

British Mathematical Colloquium, St Andrews, June 2018

Abstract

In the 1920s, R. A. Fisher, at Rothamsted Experimental Station in Harpenden, recommended Latin squares for agricultural crop experiments. At about the same time, Jerzy Neyman developed the same idea during his doctoral study at the University of Warsaw. However, there is evidence of their much earlier use in experiments.

Euler had made his famous conjecture about Graeco-Latin squares in 1782. There was a spectacular refutation in 1960.

I shall say something about the different uses of Latin squares in designed experiments. This needs methods of construction, of counting, and of randomization.

Fisher and Neyman had a famous falling out over Latin squares in 1935 when Neyman proved that use of Latin squares in experiments gives biased results. A six-week international workshop in Boulder, Colorado in 1957 resolved this, but the misunderstanding surfaced again in a Statistics paper published in 2017.

What is a Latin square?

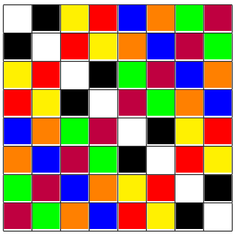
Definition

Let n be a positive integer.

A Latin square of order n is an $n \times n$ array of cells in which n symbols are placed, one per cell, in such a way that each symbol occurs once in each row and once in each column.

The symbols may be letters, numbers, colours, . . .

A Latin square of order 8



A Latin square of order 6

E	*B*	*F*	*A*	*C*	*D*
B	*C*	*D*	*E*	*F*	*A*
A	*E*	*C*	*B*	*D*	*F*
F	*D*	*E*	*C*	*A*	*B*
D	*A*	*B*	*F*	*E*	*C*
C	*F*	*A*	*D*	*B*	*E*

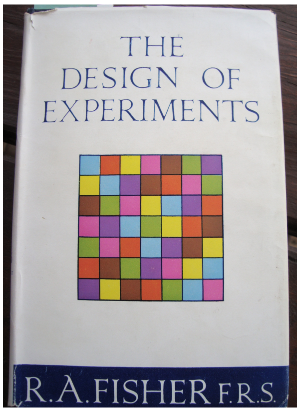
A stained glass window in Caius College, Cambridge

photograph by
J. P. Morgan

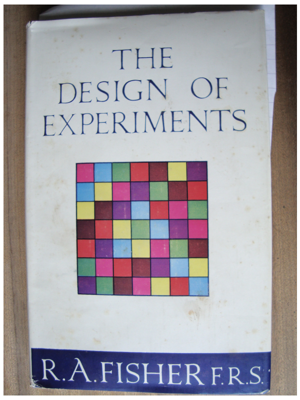
And on the opposite side of the hall

R. A. Fisher promoted the use of Latin squares in experiments while at Rothamsted (1919–1933) and his 1935 book *The Design of Experiments*.

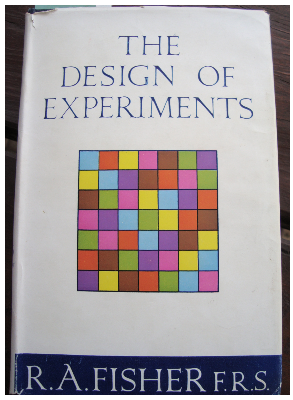
Stained glass window; book cover; INI logo



Latin squares on book covers

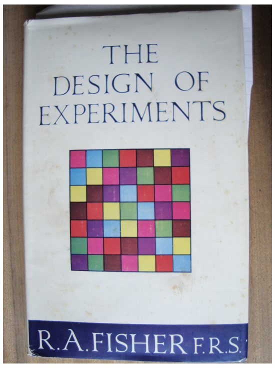


6th edition



7th edition

A Latin square of order 7



This Latin square was on the cover of the first edition of *The Design of Experiments*.

Why this one?
It does not appear in the book. It does not match any known experiment designed by Fisher.

Why is it called 'Latin'?

What are Latin squares used for?

Agricultural field trials, with rows and columns corresponding to actual rows and columns on the ground (possibly the width of rows is different from the width of columns).

"…on any given field agricultural operations, at least for centuries, have followed one of two directions, which are usually those of the rows and columns; consequently streaks of fertility, weed infestation, etc., do, in fact, occur predominantly in those two directions."

R. A. Fisher,
letter to H. Jeffreys,
30 May 1938
(selected correspondence edited by J. H. Bennett)

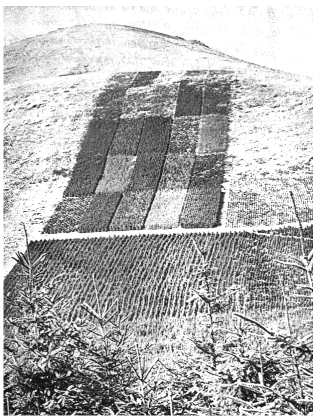
This assumption is dubious for field trials in Australia.

An experiment on potatoes at Ely in 1932

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

Treatment	A	B	C	D	E	F
Extra nitrogen	0	0	0	1	1	1
Extra phosphate	0	1	2	0	1	2

A forestry experiment



Experiment on a hillside near Beddgelert Forest, designed by Fisher and laid out in 1929

©The Forestry Commission

Other sorts of rows and columns: animals

An experiment on 16 sheep carried out by François Cretté de Palluel, reported in *Annals of Agriculture* in 1790. They were fattened on the given diet, and slaughtered on the date shown.

slaughter date	Breed: Ile de France	Beauce	Champagne	Picardy
20 Feb	potatoes	turnips	beets	oats & peas
20 Mar	turnips	beets	oats & peas	potatoes
20 Apr	beets	oats & peas	potatoes	turnips
20 May	oats & peas	potatoes	turnips	beets

Other sorts of rows and columns: plants in pots

An experiment where treatments can be applied to individual leaves of plants in pots.

height	plant 1	2	3	4
1	A	B	C	D
2	B	A	D	C
3	C	D	A	B
4	D	C	B	A

Graeco-Latin squares

A	B	C
C	A	B
B	C	A

α	β	γ
β	γ	α
γ	α	β

When the two Latin squares are superposed,
each Latin letter occurs exactly once with each Greek letter.

$A\ \alpha$	$B\ \beta$	$C\ \gamma$
$C\ \beta$	$A\ \gamma$	$B\ \alpha$
$B\ \gamma$	$C\ \alpha$	$A\ \beta$

Euler called such a superposition a 'Graeco-Latin square'. The name 'Latin square' seems to be a back-formation from this.

Pairs of orthogonal Latin squares

Definition
A pair of Latin squares of order n are orthogonal to each other if, when they are superposed, each letter of one occurs exactly once with each letter of the other.

We have just seen a pair of orthogonal Latin squares of order 3.

Mutually orthogonal Latin squares

Definition
A collection of Latin squares of the same order is mutually orthogonal if every pair is orthogonal.

Example ($n = 4$)

$A\alpha 1$	$B\beta 2$	$C\gamma 3$	$D\delta 4$
$B\gamma 4$	$A\delta 3$	$D\alpha 2$	$C\beta 1$
$C\delta 2$	$D\gamma 1$	$A\beta 4$	$B\alpha 3$
$D\beta 3$	$C\alpha 4$	$B\delta 1$	$A\gamma 2$

Theorem
If there exist k mutually orthogonal Latin squares $L_1, \ldots, L_k$ of order n, then $k \leq n - 1$.

When is the maximum achieved?

Theorem
If n is a power of a prime number then there exist $n - 1$ mutually orthogonal Latin squares of order n.

For example, $n = 2, 3, 4, 5, 7, 8, 9, 11, 13, \ldots$.

The standard construction uses a finite field of order n.

R. A. Fisher and F. Yates: *Statistical Tables for Biological, Agricultural and Medical Research*. Edinburgh, Oliver and Boyd, 1938.
This book gives a set of $n - 1$ MOLS for $n = 3, 4, 5, 7, 8$ and 9. The set of order 9 is not made by the usual finite-field construction, and it is not known how Fisher and Yates obtained this.

An industrial experiment using MOLS

L. C. H. Tippett: Applications of statistical methods to the control of quality in industrial production. Manchester Statistical Society (1934). (Cited by Fisher, 1935)

A cotton mill has 5 spindles, each made of 4 components. Why is one spindle producing defective weft?

Period	i	ii	iiii	iv	v
1	$A\alpha 1$	$B\beta 2$	$C\gamma 3$	$D\delta 4$	$E\varepsilon 5$
2	$E\delta 3$	$A\varepsilon 4$	$B\alpha 5$	$C\beta 1$	$D\gamma 2$
3	$D\beta 5$	$E\gamma 1$	$A\delta 2$	$B\varepsilon 3$	$C\alpha 4$
4	$C\varepsilon 2$	$D\alpha 3$	$E\beta 4$	$A\gamma 5$	$B\delta 1$
5	$B\gamma 4$	$C\delta 5$	$D\varepsilon 1$	$E\alpha 2$	$A\beta 3$

1st component	2nd component	3rd component	4th component
i–v	A–E	α–ε	1–5

How to randomize? I

R. A. Fisher: The arrangement of field experiments. *Journal of the Ministry of Agriculture*, **33** (1926), 503–513.

> Systematic arrangements in a square . . . have been used previously for variety trials in, for example, Ireland and Denmark; but the term "Latin square" should not be applied to any such systematic arrangements. The problem of the Latin Square, from which the name was borrowed, as formulated by Euler, consists in the enumeration of *every possible* arrangement, subject to the conditions that each row and each column shall contain one plot of each variety. Consequently, the term Latin Square should only be applied to a process of randomization by which one is selected at random out of the total number of Latin Squares possible, . . .

How many different Latin squares of order n are there?

Are these two Latin squares the same?

A	B	C
C	A	B
B	C	A

1	2	3
3	1	2
2	3	1

To answer this question, we will have to insist that all the Latin squares use the same symbols, such as $1, 2, \ldots, n$.

Reduced Latin squares, and equivalence

Definition
A Latin square is reduced if the symbols in the first row and first column are $1, 2, \ldots, n$ in natural order.

Definition
Latin squares L and M are equivalent if there is a permutation f of the rows, a permutation g of the columns and permutation h of the symbols such that

$$\begin{array}{llllllll} \text{symbol} & s & \text{is in row} & r & \text{and column} & c & \text{of} & L \\ & & & \Longleftrightarrow & & & & \\ \text{symbol} & h(s) & \text{is in row} & f(r) & \text{and column} & g(c) & \text{of} & M. \end{array}$$

Theorem
If there are m reduced squares in an equivalence class of Latin squares of order n, then the total number of Latin squares in the equivalence class is $m \times n! \times (n-1)!$.

Numbers of reduced Latin squares

order	cyclic	non-cyclic group	non-group	all	equivalence classes
2	1	0	0	1	1
3	1	0	0	1	1
4	3	1	0	4	2
5	6	0	50	56	2
6	60	80	9268	9408	22
7	120	0	16941960	16942080	564
8	1260	1500	$> 10^{12}$	$> 10^{12}$	1676267
9	6720	840	$> 10^{15}$	$> 10^{15}$	$> 10^{12}$
10	90720	36288	$> 10^{25}$	$> 10^{25}$	$> 10^{18}$
11	36288	0	$> 10^{34}$	$> 10^{34}$	$> 10^{26}$

6: Frolov, 1890; Tarry, 1900; Fisher and Yates, 1934
7: Frolov (badly wrong); Norton, 1939 (omitted one species); Sade, 1948; Saxena, 1951
8: Wells, 1967 9: Baumel and Rothstein, 1975
10: McKay and Rogoyski, 1995 11: McKay and Wanless, 2005

How to randomize? II

R. A. Fisher: *Statistical Methods for Research Workers*. Edinburgh, Oliver and Boyd, 1925.

F. Yates: The formation of Latin squares for use in field experiments. *Empire Journal of Experimental Agriculture*, **1** (1933), 235–244.

R. A. Fisher: *The Design of Experiments*. Edinburgh, Oliver and Boyd, 1935.

These three all argued that randomization should ensure validity by eliminating bias in the estimation of the difference between the effect of any two treatments, and in the estimation of the variance of the foregoing estimator. This assumes that the data analysis allows for the effects of rows and columns.

Valid randomization

Random choice of a Latin square from a given set $\mathcal{L}$ of Latin squares or order n is valid if

- every cell in the square is equally likely to have each letter (this ensures lack of bias in the estimation of the difference between treatment effects)
- every ordered pair of cells in different rows and columns has probability $1/n(n-1)$ of having the same specified letter, and probability $(n-2)/n(n-1)^2$ of having each ordered pair of distinct letters (this ensures lack of bias in the estimation of the variance).

Some methods of valid randomization

1. Permute rows by a random permutation and permute columns by an independently chosen random permutation (a.k.a. randomize rows and columns)— now the standard method.
2. Use any doubly transitive group in the above, rather than the whole symmetric group S_n (Grundy and Healy, 1950; Bailey, 1983).
3. Choose a Latin square at random from a complete set of mutually orthogonal Latin squares, and then randomize letters (Preece, Bailey and Patterson, 1978, following a 1935 remark of Fisher's when discussing a paper of Neyman).

Back to pairs of orthogonal Latin squares

Question (Euler, 1782)

For which values of n does there exist a pair of orthogonal Latin squares of order n?

Theorem

If n is odd, or if n is divisible by 4, then there is a pair of orthogonal Latin squares of order n.

Proof.

(i) If n is odd, the Latin squares with entries in (i, j) defined by $i + j$ and $i + 2j$ modulo n are mutually orthogonal.

(ii) If $n = 4$ or $n = 8$ such a pair of squares can be constructed from a finite field.

(iii) If L_1 is orthogonal to L_2, where L_1 and L_2 have order n, and M_1 is orthogonal to M_2, where M_1 and M_2 have order m, then a product construction gives squares $L_1 \otimes M_1$ orthogonal to $L_2 \otimes M_2$, where these have order nm.

Euler's conjecture

Conjecture

If n is even but not divisible by 4, then there is no pair of orthogonal Latin squares of order n.

This is true when $n = 2$, because the two letters on the main diagonal must be the same.

Euler could not find a pair of orthogonal Latin squares of order 6, or 10, or ....

Euler's conjecture: order 6

Denés and Keedwell (1974) and Klyve and Stemkovski (2006) discovered that on 10 August 1842, Heinrich Schumacher, the astronomer in Altona, Germany, wrote a letter to Gauß, telling him that his assistant, Thomas Clausen, had proved that there is no pair of orthogonal Latin squares of order 6.

He said that Clausen divided Latin squares of order 6 into 17 families, and did an exhaustive search within each family.

So had Clausen enumerated the Latin squares of order 6? This would pre-date Frolov (1890).

No written record of this proof remains.

Theorem (Tarry, 1900)

There is no pair of orthogonal Latin squares of order 6.

Proof.

Exhaustive enumeration by hand, after dividing Latin squares of order 6 into 17 families. □

The end of the conjecture

Theorem (Bose and Shrikhande, 1959)

There is a pair of orthogonal Latin squares of order 22.

Theorem (Parker, 1959)

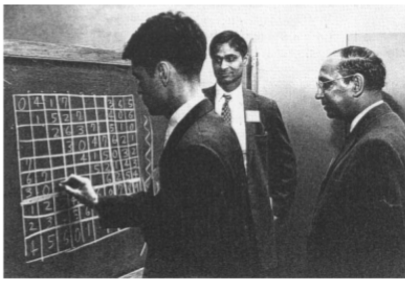
If $n = (3q - 1)/2$ and q is a power of an odd prime and $q - 3$ is divisible by 4, then there is a pair of orthogonal Latin squares of order n. In particular, there are pairs of orthogonal Latin squares of orders 10, 34, 46 and 70.

Theorem (Bose, Shrikhande and Parker, 1960)

If n is not equal to 2 or 6, then there exists a pair of orthogonal Latin squares of order n.

New York Times, 16 April 1959

Major Mathematical Conjecture Propounded 177 Years Ago Is Disproved



(Copied from *The history of latin squares* by Lars Døvling Andersen, 2007)

Some problems with Fisher's exposition

Fisher was rather authoritarian about his work.
(Ironically, he may have inadvertently mimicked Karl Pearson.)
He liked to lay down the law before the law was properly formulated and understood. But

- he rarely wrote down explicit formulae for his assumptions or methods
  (Frank Yates, his junior colleague then long-term successor at Rothamsted, did this very clearly, apparently with Fisher's blessing);
- some of his eye-catching early examples were inconsistent with his later developments
  (the lady tasting tea, and comments on an experiment of Darwin's (both in *Design of Experiments*, 1935)
  led to the randomization test,
  which he explicitly recanted in the 7th edition in 1960).

Explicit assumptions

Let $Y_\omega(i)$ be the response on plot ω ($\omega = 1, \ldots, N$) when treatment i is applied to ω.

Fisher's model is $\quad Y_\omega(i) = \tau_i + Z_\omega, \quad$ where

- τ_i depends only on treatment i,
  and we want to estimate differences like $\tau_1 - \tau_2$;
- Z_ω depends only on plot ω, and can include effects of rows and columns as well as other variability.

The additive model allows conclusions from the data analysis to be extrapolated to other plots outwith the experiment.

The joint distribution of $Z_1, \ldots, Z_N$ is partly determined by the method of randomization.

Neyman (1923, in Polish) does not assume a model for $Y_\omega(i)$, and seeks to estimate differences like

$$\frac{1}{N}\left[\sum_{\omega=1}^{N} Y_\omega(1) - \sum_{\omega=1}^{N} Y_\omega(2)\right].$$

Conclusions cannot be extrapolated.

The Fisher–Neyman row

Neyman read a paper on *Statistical problems in agricultural experimentation* to the Royal Statistical Society in 1935.
In this, he claimed to have proved that any experiment designed as a Latin square gives biased results
(in the sense that the expectation of the estimator is not equal to the true value being estimated).

Fisher responded furiously in the official discussion, but without pointing out the different underlying assumptions.

Neyman moved to the USA, where Wilk and Kempthorne (ex-Rothamsted) developed his argument further in 1957.

IMS Summer Institute

Later in 1957, Oscar Kempthorne chaired a six-week IMS Summer Institute on the topic at Boulder, Colorado.

David Cox attended this; as a result, he published a paper in 1958 explaining the misunderstanding and arguing that Fisher had been correct to state that there is no bias in a conventional Latin-square experiment. He also explained the additive assumption very clearly in his 1958 book *Planning of Experiments*.

A few years ago, Cox told me that he and Kempthorne had had really friendly discussions during the workshop.

In later years, Kempthorne (who could be as rude as Fisher in writing but as nice as pie in person) also used the additive model. In a 1975 paper he went so far as to say that Neyman's null hypothesis (that $\sum_\omega Y_\omega(i)$ is the same for every treatment i) "is not scientifically relevant".

So where are we now on this issue?

In 2017, Peng Ding published a paper in *Statistical Science*.
He claimed that Fisher's approach was to test whether $Y_\omega(i) = Y_\omega(j)$ for all i and j, even though there is no such notation in Fisher's work.
He rederived a paradox noted by George Barnard in 1955.
He ignored the IMS Summer Institute and the later papers by Kempthorne.

I was invited to contribute to the written discussion, and did so gladly and forthrightly.

Deng's response concluded

> *... as an assistant professor in the department founded by Neyman, I feel obligated to use it to continue the Neyman tradition.*