Designs with more than one blocking system: the conflict between valid randomization and latinization

R. A. Bailey
University of St Andrews

EUCARPIA, Edinburgh, September 2025

Abstract I

In experimental designs with more than one blocking system, there are (at least) two different approaches.

Abstract I

In experimental designs with more than one blocking system,
there are (at least) two different approaches.
One goes back to the work of Frank Yates and John Nelder,
who both worked at Rothamsted Experimental Station
(but at different times).

Abstract I

In experimental designs with more than one blocking system, there are (at least) two different approaches.
One goes back to the work of Frank Yates and John Nelder, who both worked at Rothamsted Experimental Station (but at different times).
They took the view that if Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells must also be included as a blocking system.

Abstract I

In experimental designs with more than one blocking system, there are (at least) two different approaches.
One goes back to the work of Frank Yates and John Nelder, who both worked at Rothamsted Experimental Station (but at different times).
They took the view that if Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells must also be included as a blocking system.
Moreover, they did not regard the variability between different plots as given by a simple equation.

Abstract I

In experimental designs with more than one blocking system, there are (at least) two different approaches.
One goes back to the work of Frank Yates and John Nelder, who both worked at Rothamsted Experimental Station (but at different times).
They took the view that if Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells must also be included as a blocking system.
Moreover, they did not regard the variability between different plots as given by a simple equation. They said that a method of randomization was valid if, when averaged over the outcomes of all possible randomizations, the expectation of the mean square for a given treatment effect is the same as the expectation of the mean square for the residual that it will be compared with, in the case that that treatment effect is zero.

Abstract II

The other approach, called latinization, seems to have begun with Boyd Harshbarger and Lyle Davis.

Abstract II

The other approach, called latinization, seems to have begun with Boyd Harshbarger and Lyle Davis.
The linear model is written in the standard form, with some factors having fixed effects and others random effects.

Abstract II

The other approach, called latinization, seems to have begun with Boyd Harshbarger and Lyle Davis.
The linear model is written in the standard form, with some factors having fixed effects and others random effects.
There is no restriction on the relationship between pairs of blocking systems.

Abstract II

The other approach, called latinization, seems to have begun with Boyd Harshbarger and Lyle Davis.

The linear model is written in the standard form, with some factors having fixed effects and others random effects.

There is no restriction on the relationship between pairs of blocking systems.

However, in each system, the number of times that a treatment occurs in any block differs by no more than one from the number times that any other treatment occurs in that block.

Abstract II

The other approach, called latinization, seems to have begun with Boyd Harshbarger and Lyle Davis.

The linear model is written in the standard form, with some factors having fixed effects and others random effects.

There is no restriction on the relationship between pairs of blocking systems.

However, in each system, the number of times that a treatment occurs in any block differs by no more than one from the number times that any other treatment occurs in that block.

I will try to explain the two approaches, and then give examples where they differ.

Introduction to Design of Experiments: Part 1

We start with 2 simple situations.

Introduction to Design of Experiments: Part 1

We start with 2 simple situations.

1. Completely randomized design

There are v treatments and vr plots.

Introduction to Design of Experiments: Part 1

We start with 2 simple situations.

1. Completely randomized design

There are v treatments and vr plots.
Allocate the treatments systematically to the plots, each treatment on r plots.

Introduction to Design of Experiments: Part 1

We start with 2 simple situations.

1. Completely randomized design

There are v treatments and vr plots.
Allocate the treatments systematically to the plots, each treatment on r plots.
Choose a random permutation from the symmetric group S_{vr} and apply it to the sytematic design. Use this for the experiment.

We start with 2 simple situations.

1. Completely randomized design

There are v treatments and vr plots.
Allocate the treatments systematically to the plots, each treatment on r plots.
Choose a random permutation from the symmetric group S_{vr} and apply it to the sytematic design. Use this for the experiment.

Conduct the experiment, obtaining one measurement on each plot.

Introduction to Design of Experiments: Part 1

We start with 2 simple situations.

1. Completely randomized design

There are v treatments and vr plots.
Allocate the treatments systematically to the plots, each treatment on r plots.
Choose a random permutation from the symmetric group S_{vr} and apply it to the sytematic design. Use this for the experiment.

Conduct the experiment, obtaining one measurement on each plot.
Estimate each treatment effect by the mean yield on the plots with that treatment.

We start with 2 simple situations.

1. Completely randomized design

There are v treatments and vr plots.
Allocate the treatments systematically to the plots, each treatment on r plots.
Choose a random permutation from the symmetric group S_{vr} and apply it to the sytematic design. Use this for the experiment.

Conduct the experiment, obtaining one measurement on each plot.
Estimate each treatment effect by the mean yield on the plots with that treatment.
Use a function of the residual mean square (using data orthogonal to the treatment subspace) to estimate the variance of those estimates.

2. Complete-block design

There are v treatments.
There are vr plots, which are arranged in r blocks of size v.

2. Complete-block design

There are v treatments.
There are vr plots, which are arranged in r blocks of size v.
Within each block, allocate one treatment to each plot.

2. Complete-block design

There are v treatments.
There are vr plots, which are arranged in r blocks of size v.
Within each block, allocate one treatment to each plot.
Within each block independently, choose a random permutation from the symmetric group S_v and apply it to the plots in that block. Use the outcome for the experiment.

2. Complete-block design

There are v treatments.
There are vr plots, which are arranged in r blocks of size v.
Within each block, allocate one treatment to each plot.
Within each block independently, choose a random permutation from the symmetric group S_v and apply it to the plots in that block. Use the outcome for the experiment.

Conduct the experiment, obtaining one measurement on each plot.

2. Complete-block design

There are v treatments.
There are vr plots, which are arranged in r blocks of size v.
Within each block, allocate one treatment to each plot.
Within each block independently, choose a random permutation from the symmetric group S_v and apply it to the plots in that block. Use the outcome for the experiment.

Conduct the experiment, obtaining one measurement on each plot.
Estimate each treatment difference by the difference between the mean yields on the plots with those two treatments.

2. Complete-block design

There are v treatments.
There are vr plots, which are arranged in r blocks of size v.
Within each block, allocate one treatment to each plot.
Within each block independently, choose a random permutation from the symmetric group S_v and apply it to the plots in that block. Use the outcome for the experiment.

Conduct the experiment, obtaining one measurement on each plot.
Estimate each treatment difference by the difference between the mean yields on the plots with those two treatments.
Use a function of the residual mean square (using data orthogonal to the treatment subspace and orthogonal to the block subspace) to estimate the variance of those estimates.

More on Complete-Block designs

When analysing the data from such an experiment,
some people forget to include the blocks in the analysis.

More on Complete-Block designs

When analysing the data from such an experiment,
some people forget to include the blocks in the analysis.

Some other people do not forget, but first perform a significance test to see whether there is a significant difference in block effects, and then write something like

> *Blocks were found to have no statistically significant effects so we removed them from all subsequent analysis.*

More on Complete-Block designs

When analysing the data from such an experiment, some people forget to include the blocks in the analysis.

Some other people do not forget, but first perform a significance test to see whether there is a significant difference in block effects, and then write something like

> *Blocks were found to have no statistically significant effects so we removed them from all subsequent analysis.*

Since the blocks mean square may well be larger than the residual mean square (even if it is not larger enough to be "significant"), this procedure inflates the residual mean square and so make it harder to detect real differences between treatments.

More on Complete-Block designs

When analysing the data from such an experiment, some people forget to include the blocks in the analysis.

Some other people do not forget, but first perform a significance test to see whether there is a significant difference in block effects, and then write something like

> *Blocks were found to have no statistically significant effects so we removed them from all subsequent analysis.*

Since the blocks mean square may well be larger than the residual mean square (even if it is not larger enough to be "significant"), this procedure inflates the residual mean square and so make it harder to detect real differences between treatments.

R. A. Fisher pointed out this mistake in his 1925 book *Statistical Methods for Research Workers*.

More on Complete-Block designs

When analysing the data from such an experiment,
some people forget to include the blocks in the analysis.

Some other people do not forget, but first perform a significance test to see whether there is a significant difference in block effects, and then write something like

> *Blocks were found to have no statistically significant effects so we removed them from all subsequent analysis.*

Since the blocks mean square may well be larger than the residual mean square (even if it is not larger enough to be "significant"), this procedure inflates the residual mean square and so make it harder to detect real differences between treatments.

R. A. Fisher pointed out this mistake in his 1925 book *Statistical Methods for Research Workers*.

Unfortunately, this mistake is still made today,
especially in clinical trials.

What about Latin squares? Definition

A Latin square of order v has v treatments in a $v \times v$ square array, with each treatment occuring once in each row and once in each column.

What about Latin squares? Definition

A Latin square of order v has v treatments in a $v \times v$ square array, with each treatment occuring once in each row and once in each column. Here is one with $v = 5$.

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

What about Latin squares? Definition

A Latin square of order v has v treatments in a $v \times v$ square array, with each treatment occuring once in each row and once in each column. Here is one with $v = 5$.

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

With Completely Randomized Designs and with Complete Block Designs, the method of randomization shown before gives all possible designs with that structure.

What about Latin squares? Definition

A Latin square of order v has v treatments in a $v \times v$ square array, with each treatment occuring once in each row and once in each column. Here is one with $v = 5$.

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

With Completely Randomized Designs and with Complete Block Designs, the method of randomization shown before gives all possible designs with that structure.
Maybe that is why Ronald Fisher thought that the proper way to randomize a Latin square was to enumerate all Latin squares of order v and choose one at random?

What about Latin squares? Definition

A Latin square of order v has v treatments in a $v \times v$ square array, with each treatment occuring once in each row and once in each column. Here is one with $v = 5$.

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

With Completely Randomized Designs and with Complete Block Designs, the method of randomization shown before gives all possible designs with that structure.
Maybe that is why Ronald Fisher thought that the proper way to randomize a Latin square was to enumerate all Latin squares of order v and choose one at random?
(Or maybe he just loved Latin squares so much that he wanted an excuse to do that enumeration?)

What about Latin squares? How to randomize?

In 1931, Frank Yates became Fisher's colleague at Rothamsted Experimental Station, where he spent the rest of his career.

What about Latin squares? How to randomize?

In 1931, Frank Yates became Fisher's colleague at Rothamsted Experimental Station, where he spent the rest of his career.

In his 1933 paper on "The formation of Latin squares for use in field experiments", he showed that there is no need to choose a Latin square at random from all those of that size.

What about Latin squares? How to randomize?

In 1931, Frank Yates became Fisher's colleague at Rothamsted Experimental Station, where he spent the rest of his career.

In his 1933 paper on "The formation of Latin squares for use in field experiments", he showed that there is no need to choose a Latin square at random from all those of that size.

He showed that it is sufficient to start with one Latin square of order v, then randomize rows using a random permutation from S_v, and then randomize columns by using an independently chosen random permutation from S_v.

More systems of blocks

A semi-Latin square is an extension of the idea of Latin square. Now there are n^2k plots, which are arranged in a $n \times n$ array of cells, where each cell has k plots. There are nk treatments, each allocated to one plot in each row and one plot in each column.

More systems of blocks

A semi-Latin square is an extension of the idea of Latin square. Now there are n^2k plots, which are arranged in a $n \times n$ array of cells, where each cell has k plots. There are nk treatments, each allocated to one plot in each row and one plot in each column. Here is one with $n = 4$ and $k = 3$.

A E I	B F J	C G K	D H L
B H K	A G L	D F I	C E J
C F L	D E K	A H J	B G I
D G J	C H I	B E L	A F K

More systems of blocks

A semi-Latin square is an extension of the idea of Latin square. Now there are n^2k plots, which are arranged in a $n \times n$ array of cells, where each cell has k plots. There are nk treatments, each allocated to one plot in each row and one plot in each column. Here is one with $n = 4$ and $k = 3$.

A E I	B F J	C G K	D H L
B H K	A G L	D F I	C E J
C F L	D E K	A H J	B G I
D G J	C H I	B E L	A F K

In his 1935 paper on “Complex experiments”, Yates said that this arrangement “stands condemned”, because he saw data being analysed without including cells as a blocking factor.

More systems of blocks

A semi-Latin square is an extension of the idea of Latin square. Now there are n^2k plots, which are arranged in a $n \times n$ array of cells, where each cell has k plots. There are nk treatments, each allocated to one plot in each row and one plot in each column. Here is one with $n = 4$ and $k = 3$.

$A \quad E \quad I$	$B \quad F \quad J$	$C \quad G \quad K$	$D \quad H \quad L$
$B \quad H \quad K$	$A \quad G \quad L$	$D \quad F \quad I$	$C \quad E \quad J$
$C \quad F \quad L$	$D \quad E \quad K$	$A \quad H \quad J$	$B \quad G \quad I$
$D \quad G \quad J$	$C \quad H \quad I$	$B \quad E \quad L$	$A \quad F \quad K$

In his 1935 paper on "Complex experiments", Yates said that this arrangement "stands condemned", because he saw data being analysed without including cells as a blocking factor. They essentially form a system of incomplete blocks (which means that each cell contains less than all the treatments).

More systems of blocks

A semi-Latin square is an extension of the idea of Latin square. Now there are n^2k plots, which are arranged in a $n \times n$ array of cells, where each cell has k plots. There are nk treatments, each allocated to one plot in each row and one plot in each column. Here is one with $n = 4$ and $k = 3$.

A E I	B F J	C G K	D H L
B H K	A G L	D F I	C E J
C F L	D E K	A H J	B G I
D G J	C H I	B E L	A F K

In his 1935 paper on "Complex experiments", Yates said that this arrangement "stands condemned", because he saw data being analysed without including cells as a blocking factor. They essentially form a system of incomplete blocks (which means that each cell contains less than all the treatments). Once people started to include the cells in the data analysis, he became enthusiastic about of the use of semi-Latin squares.

Orthogonal block structures

John Nelder took over from Frank Yates as head of the Statistics Department at Rothamsted from 1968 to 1984.

Orthogonal block structures

John Nelder took over from Frank Yates as head of the Statistics Department at Rothamsted from 1968 to 1984.

In two papers in 1965, he had introduced the general family of orthogonal block structures, which includes all those of Yates.

Orthogonal block structures

John Nelder took over from Frank Yates as head of the Statistics Department at Rothamsted from 1968 to 1984.

In two papers in 1965, he had introduced the general family of orthogonal block structures, which includes all those of Yates.

They all have the following properties.

1. If there are two systems of blocks, say Rows and Columns, then the system whose blocks are all intersections of a Row with a Column is also included.

Orthogonal block structures

John Nelder took over from Frank Yates as head of the Statistics Department at Rothamsted from 1968 to 1984.

In two papers in 1965, he had introduced the general family of orthogonal block structures, which includes all those of Yates.

They all have the following properties.

1. If there are two systems of blocks, say Rows and Columns, then the system whose blocks are all intersections of a Row with a Column is also included.
2. Each system of blocks has all its blocks of the same size.

Orthogonal block structures

John Nelder took over from Frank Yates as head of the Statistics Department at Rothamsted from 1968 to 1984.

In two papers in 1965, he had introduced the general family of orthogonal block structures, which includes all those of Yates.

They all have the following properties.

1. If there are two systems of blocks, say Rows and Columns, then the system whose blocks are all intersections of a Row with a Column is also included.
2. Each system of blocks has all its blocks of the same size.
3. If Rows and Columns are two systems of blocks then they are orthogonal to each other.
   (I will not go into the technicalities of this definition here, but just say that it is satisfied in all of my examples.)

Valid randomization

The linear model that we assume here is not the simple text-book one. We assume that for each treatment i there is a (usually fixed) effect τ_i and that there is a number (maybe fixed, maybe random) ζ_ω such that if treatment i is applied to plot ω then the response is $\tau_i + \zeta_\omega$.

Valid randomization

The linear model that we assume here is not the simple text-book one. We assume that for each treatment i there is a (usually fixed) effect τ_i and that there is a number (maybe fixed, maybe random) ζ_ω such that if treatment i is applied to plot ω then the response is $\tau_i + \zeta_\omega$.

Following a suggestion of Frank Yates, his Rothamsted colleagues (Patrick) Mike Grundy and Michael Healy defined a method of randomization to be strongly valid if, for any set of linear combinations of (orthogonal) treatment comparisons, when averaged over all possible outcomes of randomization, the average mean square for that treatment subspace is equal to the average mean square for error in the case that τ_i is constant for all i.

Valid randomization

The linear model that we assume here is not the simple text-book one. We assume that for each treatment i there is a (usually fixed) effect τ_i and that there is a number (maybe fixed, maybe random) ζ_ω such that if treatment i is applied to plot ω then the response is $\tau_i + \zeta_\omega$.

Following a suggestion of Frank Yates, his Rothamsted colleagues (Patrick) Mike Grundy and Michael Healy defined a method of randomization to be strongly valid if, for any set of linear combinations of (orthogonal) treatment comparisons, when averaged over all possible outcomes of randomization, the average mean square for that treatment subspace is equal to the average mean square for error in the case that τ_i is constant for all i.

Of course, this definition assumes that we are using the appropriate error mean square.

Grundy and Healy, continued

Grundy and Healy were concerned about a $2 \times 2 \times 2 \times 2 \times 4$ fractional factorial experiment in an 8×8 square which had been carried out at Rothamsted a few years earlier. It had not been noticed at the time that one of the 2-level factors had one of its levels in the 4×4 subsquares of contiguous plots in the top left and the bottom right.

Grundy and Healy, continued

Grundy and Healy were concerned about a $2 \times 2 \times 2 \times 2 \times 4$ fractional factorial experiment in an 8×8 square which had been carried out at Rothamsted a few years earlier. It had not been noticed at the time that one of the 2-level factors had one of its levels in the 4×4 subsquares of contiguous plots in the top left and the bottom right.

They wanted to find a valid procedure for an 8×8 Latin square that would ensure that the outcome of this would not contain any 4×4 subsquare of contiguous plots with only 4 letters.

Grundy and Healy, continued

Grundy and Healy were concerned about a $2 \times 2 \times 2 \times 2 \times 4$ fractional factorial experiment in an 8×8 square which had been carried out at Rothamsted a few years earlier. It had not been noticed at the time that one of the 2-level factors had one of its levels in the 4×4 subsquares of contiguous plots in the top left and the bottom right.

They wanted to find a valid procedure for an 8×8 Latin square that would ensure that the outcome of this would not contain any 4×4 subsquare of contiguous plots with only 4 letters.

They showed that validity can be obtained by using a doubly transitive subgroup G of S_8.
(This means that if $\alpha \neq \gamma$ and $\beta \neq \delta$ then there is at least one permutation in G which takes α to β and γ to δ.)

Grundy and Healy, continued

Grundy and Healy were concerned about a $2 \times 2 \times 2 \times 2 \times 4$ fractional factorial experiment in an 8×8 square which had been carried out at Rothamsted a few years earlier. It had not been noticed at the time that one of the 2-level factors had one of its levels in the 4×4 subsquares of contiguous plots in the top left and the bottom right.

They wanted to find a valid procedure for an 8×8 Latin square that would ensure that the outcome of this would not contain any 4×4 subsquare of contiguous plots with only 4 letters.

They showed that validity can be obtained by using a doubly transitive subgroup G of S_8.
(This means that if $\alpha \neq \gamma$ and $\beta \neq \delta$ then there is at least one permutation in G which takes α to β and γ to δ.)
Then randomize rows using a permutation chosen at random from G and independently randomize columns by choosing a permutation at random from G.

Grundy and Healy, continued

Grundy and Healy were concerned about a $2 \times 2 \times 2 \times 2 \times 4$ fractional factorial experiment in an 8×8 square which had been carried out at Rothamsted a few years earlier. It had not been noticed at the time that one of the 2-level factors had one of its levels in the 4×4 subsquares of contiguous plots in the top left and the bottom right.

They wanted to find a valid procedure for an 8×8 Latin square that would ensure that the outcome of this would not contain any 4×4 subsquare of contiguous plots with only 4 letters.

They showed that validity can be obtained by using a doubly transitive subgroup G of S_8.
(This means that if $\alpha \neq \gamma$ and $\beta \neq \delta$ then there is at least one permutation in G which takes α to β and γ to δ.)
Then randomize rows using a permutation chosen at random from G and independently randomize columns by choosing a permutation at random from G.

They succeeded in finding a suitable starting array.

Youden's approach

With no knowledge of Grundy and Healy's approach, Youden gave a different method of valid randomization. He wrote this in 1956, but it was not published until 1972.

Youden's approach

With no knowledge of Grundy and Healy's approach, Youden gave a different method of valid randomization. He wrote this in 1956, but it was not published until 1972. Here is how his method works for a row of 9 plots with 3 treatments. (This example is taken from Nelson and Bailey, 2003.)

Youden's approach

With no knowledge of Grundy and Healy's approach, Youden gave a different method of valid randomization. He wrote this in 1956, but it was not published until 1972. Here is how his method works for a row of 9 plots with 3 treatments. (This example is taken from Nelson and Bailey, 2003.)

1. Decide on some property of the design that is undesirable (for this example, there should be no 3 consecutive plots with the same treatment, nor should any treatment have all of its occurrences in the first 4 or the last 4 plots).

Youden's approach

With no knowledge of Grundy and Healy's approach, Youden gave a different method of valid randomization. He wrote this in 1956, but it was not published until 1972. Here is how his method works for a row of 9 plots with 3 treatments. (This example is taken from Nelson and Bailey, 2003.)

1. Decide on some property of the design that is undesirable (for this example, there should be no 3 consecutive plots with the same treatment, nor should any treatment have all of its occurrences in the first 4 or the last 4 plots).
2. Create a catalogue of rows of length 9 which includes no undesirable ones but has the property that, for every pair of plots, there are precisely one quarter of the rows in which they have the same treatment. (If A occurs in plot 1, then it also occurs in 2 of the remaining 8 plots, so the number of rows in the catalogue must be a multiple of 4.)

Youden's approach

With no knowledge of Grundy and Healy's approach, Youden gave a different method of valid randomization. He wrote this in 1956, but it was not published until 1972. Here is how his method works for a row of 9 plots with 3 treatments. (This example is taken from Nelson and Bailey, 2003.)

1. Decide on some property of the design that is undesirable (for this example, there should be no 3 consecutive plots with the same treatment, nor should any treatment have all of its occurrences in the first 4 or the last 4 plots).
2. Create a catalogue of rows of length 9 which includes no undesirable ones but has the property that, for every pair of plots, there are precisely one quarter of the rows in which they have the same treatment. (If A occurs in plot 1, then it also occurs in 2 of the remaining 8 plots, so the number of rows in the catalogue must be a multiple of 4.)
3. Choose one row from the catalogue at random.

Youden's approach

With no knowledge of Grundy and Healy's approach, Youden gave a different method of valid randomization. He wrote this in 1956, but it was not published until 1972. Here is how his method works for a row of 9 plots with 3 treatments. (This example is taken from Nelson and Bailey, 2003.)

1. Decide on some property of the design that is undesirable (for this example, there should be no 3 consecutive plots with the same treatment, nor should any treatment have all of its occurrences in the first 4 or the last 4 plots).
2. Create a catalogue of rows of length 9 which includes no undesirable ones but has the property that, for every pair of plots, there are precisely one quarter of the rows in which they have the same treatment. (If A occurs in plot 1, then it also occurs in 2 of the remaining 8 plots, so the number of rows in the catalogue must be a multiple of 4.)
3. Choose one row from the catalogue at random.
4. Randomly allocate the 3 actual treatments to the letters A, B, C.

Example of using Youden's approach

Choose a row at random from the list below.

$$\begin{array}{ccccccccc}
A & B & A & C & A & C & C & B & B \\
A & A & B & B & C & A & C & B & C \\
A & B & B & A & C & C & B & C & A \\
A & B & C & B & B & C & A & A & C
\end{array}$$

Example of using Youden's approach

Choose a row at random from the list below.

$$\begin{array}{ccccccccc}
A & B & A & C & A & C & C & B & B \\
A & A & B & B & C & A & C & B & C \\
A & B & B & A & C & C & B & C & A \\
A & B & C & B & B & C & A & A & C
\end{array}$$

Then randomly allocate the three actual treatments to the letters A, B, C.

Connection between block systems and valid randomization

Suppose that rows and columns form two systems of blocks.
Call the intersection of a row and a column a cell.
Suppose that each cell has k plots, where $k > 1$.

Connection between block systems and valid randomization

Suppose that rows and columns form two systems of blocks.
Call the intersection of a row and a column a cell.
Suppose that each cell has k plots, where $k > 1$.

In a 1987 paper, Bailey and Rowley proved that if the block system contains both rows and columns then there cannot be a valid method of randomization unless the block system also includes cells.

Correlation

Suppose that α and β are two different plots.

Correlation

Suppose that α and β are two different plots.

Denote by $\rho_{\alpha\beta}$ the correlation between the variables measured on those two plots.

Correlation

Suppose that α and β are two different plots.

Denote by $\rho_{\alpha\beta}$ the correlation between the variables measured on those two plots.

In the classical theory of linear models, all such correlations are non-negative.

Correlation

Suppose that α and β are two different plots.

Denote by $\rho_{\alpha\beta}$ the correlation between the variables measured on those two plots.

In the classical theory of linear models, all such correlations are non-negative.

My approach specifies when $\rho_{\alpha\beta}$ must be equal to $\rho_{\gamma\delta}$ but it does not insist that the correlations are non-negative.

Correlation

Suppose that α and β are two different plots.

Denote by $\rho_{\alpha\beta}$ the correlation between the variables measured on those two plots.

In the classical theory of linear models, all such correlations are non-negative.

My approach specifies when $\rho_{\alpha\beta}$ must be equal to $\rho_{\gamma\delta}$ but it does not insist that the correlations are non-negative.

For example, if the experiment is conducted in a greenhouse, with plants of several different varieties in each pot, then competition for resources may well lead to negative correlations in outcomes per plant in the same pot.

Summary of my approach

1. In each system of blocks used, all blocks have the same size.

Summary of my approach

1. In each system of blocks used, all blocks have the same size.
2. If any system of blocks is included in the design of the experiment, it must also be included in the data analysis.

Summary of my approach

1. In each system of blocks used, all blocks have the same size.
2. If any system of blocks is included in the design of the experiment, it must also be included in the data analysis.
3. If Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells must also be included as a blocking system (unless each cell is an individual plot).

Summary of my approach

1. In each system of blocks used, all blocks have the same size.
2. If any system of blocks is included in the design of the experiment, it must also be included in the data analysis.
3. If Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells must also be included as a blocking system (unless each cell is an individual plot).
4. Randomization must be valid for the proposed method of data analysis.

Summary of my approach

1. In each system of blocks used, all blocks have the same size.
2. If any system of blocks is included in the design of the experiment, it must also be included in the data analysis.
3. If Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells must also be included as a blocking system (unless each cell is an individual plot).
4. Randomization must be valid for the proposed method of data analysis.
5. The response for treatment i on plot ω has the form $\tau_i + \zeta_\omega$, but we do not make any assumptions about the ζ_ω before randomization.

Summary of my approach

1. In each system of blocks used, all blocks have the same size.
2. If any system of blocks is included in the design of the experiment, it must also be included in the data analysis.
3. If Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells must also be included as a blocking system (unless each cell is an individual plot).
4. Randomization must be valid for the proposed method of data analysis.
5. The response for treatment i on plot ω has the form $\tau_i + \zeta_\omega$, but we do not make any assumptions about the ζ_ω before randomization.
6. Correlations may be negative.

Summary of my approach

1. In each system of blocks used, all blocks have the same size.
2. If any system of blocks is included in the design of the experiment, it must also be included in the data analysis.
3. If Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells must also be included as a blocking system (unless each cell is an individual plot).
4. Randomization must be valid for the proposed method of data analysis.
5. The response for treatment i on plot ω has the form $\tau_i + \zeta_\omega$, but we do not make any assumptions about the ζ_ω before randomization.
6. Correlations may be negative.

Each of those aligns well with Yates's views in his comments in his book *Experimental Design: Selected Papers*.

Latinization

Harshbarger and Davis (1952) introduced what they called Latinization.

Latinization

Harshbarger and Davis (1952) introduced what they called Latinization. Their main example has 48 plots arranged in 4 rows consisting of 4 blocks of size 3. There are also 4 sets, each consisting of one block from each row. These sets are shown as squares in the diagram.

A	B	C
D	F	E
I	G	H
L	K	J

D	G	J
A	H	K
B	E	L
C	F	I

E	I	K
C	G	L
A	F	J
B	D	H

F	H	L
B	I	J
C	D	K
A	E	G

Latinization

Harshbarger and Davis (1952) introduced what they called Latinization. Their main example has 48 plots arranged in 4 rows consisting of 4 blocks of size 3. There are also 4 sets, each consisting of one block from each row. These sets are shown as squares in the diagram.

A	B	C
D	F	E
I	G	H
L	K	J

D	G	J
A	H	K
B	E	L
C	F	I

E	I	K
C	G	L
A	F	J
B	D	H

F	H	L
B	I	J
C	D	K
A	E	G

Their data analysis allows for Rows, Sets and Blocks.

Latinization

Harshbarger and Davis (1952) introduced what they called Latinization. Their main example has 48 plots arranged in 4 rows consisting of 4 blocks of size 3. There are also 4 sets, each consisting of one block from each row. These sets are shown as squares in the diagram.

A	B	C	D	G	J	E	I	K	F	H	L
D	F	E	A	H	K	C	G	L	B	I	J
I	G	H	B	E	L	A	F	J	C	D	K
L	K	J	C	F	I	B	D	H	A	E	G

Their data analysis allows for Rows, Sets and Blocks.

They did not actually define what they meant by Latinized. Perhaps they meant that each treatment occurs once in each Row and once in each Set?

Latinization

Harshbarger and Davis (1952) introduced what they called Latinization. Their main example has 48 plots arranged in 4 rows consisting of 4 blocks of size 3. There are also 4 sets, each consisting of one block from each row. These sets are shown as squares in the diagram.

A	B	C
D	F	E
I	G	H
L	K	J

D	G	J
A	H	K
B	E	L
C	F	I

E	I	K
C	G	L
A	F	J
B	D	H

F	H	L
B	I	J
C	D	K
A	E	G

Their data analysis allows for Rows, Sets and Blocks.

They did not actually define what they meant by Latinized. Perhaps they meant that each treatment occurs once in each Row and once in each Set?

Such designs are usually called semi-Latin squares today.

Possible bad outcome of randomization

Williams (1986) pointed out that one possible outcome of starting with Harshbarger and Davis's design and then randomizing Sets, Rows and then Plots within each Set–Row combination could be the one shown here.

A	B	C
D	F	E
I	G	H
L	K	J

D	G	J
A	K	H
L	B	E
I	F	C

I	K	E
L	G	C
A	F	J
D	B	H

L	F	H
I	B	J
D	K	C
A	G	E

Possible bad outcome of randomization

Williams (1986) pointed out that one possible outcome of starting with Harshbarger and Davis's design and then randomizing Sets, Rows and then Plots within each Set–Row combination could be the one shown here.

A	B	C
D	F	E
I	G	H
L	K	J

D	G	J
A	K	H
L	B	E
I	F	C

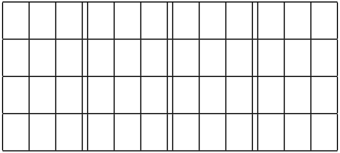
I	K	E
L	G	C
A	F	J
D	B	H

L	F	H
I	B	J
D	K	C
A	G	E

In an agricultural field trial it is likely that each column consists of contiguous plots, in which case this design is not desirable.

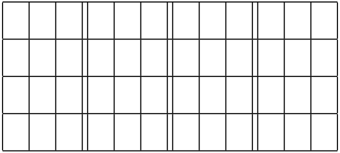
Include another blocking system

Williams (1986) proposed avoiding this problem by including Columns within Sets as another blocking system. Here is an example.



Include another blocking system

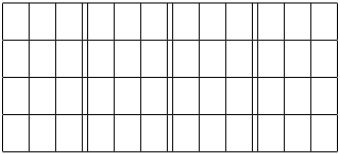
Williams (1986) proposed avoiding this problem by including Columns within Sets as another blocking system. Here is an example.



Now the data analysis has to take account of Sets, Rows, Blocks and Columns.

Include another blocking system

Williams (1986) proposed avoiding this problem by including Columns within Sets as another blocking system. Here is an example.



Now the data analysis has to take account of Sets, Rows, Blocks and Columns.

This paper also proposes a generalization of Latinization.
This is satisfied by a given system of blocks if
there is some integer λ such that each block contains each treatment either λ or $\lambda + 1$ times.

Terminology

Emlyn Williams often uses the term replicates for the blocks in one of his blocking systens.

Terminology

Emlyn Williams often uses the term replicates for the blocks in one of his blocking systens.
I prefer not to do this, for the following reasons.

Terminology

Emlyn Williams often uses the term replicates for the blocks in one of his blocking systens.

I prefer not to do this, for the following reasons.

1. I think that the blocking systems are defined before the decision is made as to how to allocate treatments to plots.

Terminology

Emlyn Williams often uses the term replicates for the blocks in one of his blocking systens.

I prefer not to do this, for the following reasons.

1. I think that the blocking systems are defined before the decision is made as to how to allocate treatments to plots.
2. In some cases, such as a Latin square, there may be more than one system of blocks (for example, Rows and Columns) where each treatment occurs once in each relevant block. We cannot call both of these “replicates”.

Extending the concept of Latinizing

John and Williams (1998) extended the concept of Latinizing to t-Latinizing, where t is an integer with $t > 1$.

Extending the concept of Latinizing

John and Williams (1998) extended the concept of Latinizing to t-Latinizing, where t is an integer with $t > 1$.
Now the r replicates form rows, each consisting of s blocks of size k, which are grouped into columns of size rk. These columns are called long blocks. Then adjacent long blocks are merged into s/t groups containing t long blocks.

Extending the concept of Latinizing

John and Williams (1998) extended the concept of Latinizing to t-Latinizing, where t is an integer with $t > 1$.

Now the r replicates form rows, each consisting of s blocks of size k, which are grouped into columns of size rk. These columns are called long blocks. Then adjacent long blocks are merged into s/t groups containing t long blocks.

Rep 1	N	K	A	V	E	M
	X	P	T	D	I	R
	C	F	Q	W	L	O
	H	J	B	U	S	G
Rep 2	T	G	H	X	D	N
	M	A	R	O	C	Q
	L	S	I	F	B	P
	U	V	J	E	K	W

Here $r = 2$, $s = 6$, $k = 4$ and $t = 2$.

Extending the concept of Latinizing

John and Williams (1998) extended the concept of Latinizing to t-Latinizing, where t is an integer with $t > 1$.
Now the r replicates form rows, each consisting of s blocks of size k, which are grouped into columns of size rk. These columns are called long blocks. Then adjacent long blocks are merged into s/t groups containing t long blocks.

Rep 1	N	K	A	V	E	M
	X	P	T	D	I	R
	C	F	Q	W	L	O
	H	J	B	U	S	G
Rep 2	T	G	H	X	D	N
	M	A	R	O	C	Q
	L	S	I	F	B	P
	U	V	J	E	K	W

Here $r = 2$, $s = 6$, $k = 4$ and $t = 2$.

The rule is that no treatment occurs more than once in any group.

Extending the concept of Latinizing, continued

Rep 1	*N*	*K*	*A*	*V*	*E*	*M*
	X	*P*	*T*	*D*	*I*	*R*
	C	*F*	*Q*	*W*	*L*	*O*
	H	*J*	*B*	*U*	*S*	*G*
Rep 2	*T*	*G*	*H*	*X*	*D*	*N*
	M	*A*	*R*	*O*	*C*	*Q*
	L	*S*	*I*	*F*	*B*	*P*
	U	*V*	*J*	*E*	*K*	*W*

Extending the concept of Latinizing, continued

Rep 1	N	K	A	V	E	M
	X	P	T	D	I	R
	C	F	Q	W	L	O
	H	J	B	U	S	G
Rep 2	T	G	H	X	D	N
	M	A	R	O	C	Q
	L	S	I	F	B	P
	U	V	J	E	K	W

The paper does not mention how Long Blocks or Groups will be used in the data analysis.

Extending the concept of Latinizing, continued

Rep 1	N	K	A	V	E	M
	X	P	T	D	I	R
	C	F	Q	W	L	O
	H	J	B	U	S	G
Rep 2	T	G	H	X	D	N
	M	A	R	O	C	Q
	L	S	I	F	B	P
	U	V	J	E	K	W

The paper does not mention how Long Blocks or Groups will be used in the data analysis.

Also, the intersections of a Replicate with a Group (one is shown in red) form another system of blocks, but this one is not mentioned in the paper.

Extending the concept of Latinizing, continued

	N	*K*	*A*	*V*	*E*	*M*
Rep 1	*X*	*P*	*T*	*D*	*I*	*R*
	C	*F*	*Q*	*W*	*L*	*O*
	H	*J*	*B*	*U*	*S*	*G*
	T	*G*	*H*	*X*	*D*	*N*
Rep 2	*M*	*A*	*R*	*O*	*C*	*Q*
	L	*S*	*I*	*F*	*B*	*P*
	U	*V*	*J*	*E*	*K*	*W*

The paper does not mention how Long Blocks or Groups will be used in the data analysis.

Also, the intersections of a Replicate with a Group (one is shown in red) form another system of blocks, but this one is not mentioned in the paper.

So it seems that there can be no valid system of randomization.

Unequal block sizes

Patterson and Williams (1976) introduced a class of resolvable incomplete-block designs which they called α-designs.

Unequal block sizes

Patterson and Williams (1976) introduced a class of resolvable incomplete-block designs which they called α-designs.

Some of these have blocks with unequal sizes.
In such cases there can be no valid randomization.

Summary of differences from my approach

1. The topic latinization seems to be rather important, but does not seem to always have exactly the same meaning.

Summary of differences from my approach

1. The topic latinization seems to be rather important, but does not seem to always have exactly the same meaning.
2. In a given system of blocks, it is not always necessary for all blocks to have the same size.

Summary of differences from my approach

1. The topic latinization seems to be rather important, but does not seem to always have exactly the same meaning.
2. In a given system of blocks, it is not always necessary for all blocks to have the same size.
3. It is not always clear whether all block systems used are included in the data analysis.

Summary of differences from my approach

1. The topic latinization seems to be rather important, but does not seem to always have exactly the same meaning.
2. In a given system of blocks, it is not always necessary for all blocks to have the same size.
3. It is not always clear whether all block systems used are included in the data analysis.
4. If Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells are not always included as a blocking system.

Summary of differences from my approach

1. The topic latinization seems to be rather important, but does not seem to always have exactly the same meaning.
2. In a given system of blocks, it is not always necessary for all blocks to have the same size.
3. It is not always clear whether all block systems used are included in the data analysis.
4. If Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells are not always included as a blocking system.
5. There is little or no discussion of what makes randomization valid.

Summary of differences from my approach

1. The topic latinization seems to be rather important, but does not seem to always have exactly the same meaning.
2. In a given system of blocks, it is not always necessary for all blocks to have the same size.
3. It is not always clear whether all block systems used are included in the data analysis.
4. If Rows and Columns are two necessary blocking systems, and each intersection of a Row and a Column is called a Cell, then Cells are not always included as a blocking system.
5. There is little or no discussion of what makes randomization valid.
6. The response for treatment i on plot ω has the form $\tau_i + \zeta_\omega$, where ζ_ω is a sum of terms depending on the plot structure.

References: Essentials of Design of Experiments

- R. A. Fisher: *Statistical Methods for Research Workers*, Oliver and Boyd, 1925.
- F. Yates: The formation of Latin squares for use in field experiments. *The Empire Journal of Experimental Agriculture*, **1**, (1933), 235–244
- F. Yates: Complex experiments. *Supplement to the Journal of the Royal Statistical Society*, **2**, (1935), 181–247.
- J. A. Nelder: The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. *Proceedings of the Royal Society of London, Series A*, **283**, (1965), 148–162.
- J. A. Nelder: The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. *Proceedings of the Royal Society of London, Series A*, **283**, (1965), 163–178.

References: Valid Randomization

- P. M. Grundy and M. J. R. Healy:
  Restricted randomization and Quasi-Latin squares, *Journal of the Royal Statistical Society, Series B*, **12**, (1950), 286–291.
- W. J. Youden: Randomization and experimentation.
  Technometrics, **14**, (1972), 13–22.
- R. A. Bailey and P. R. Nelson: Hadamard randomization: a valid restriction of random permuted blocks.
  Biometrical Journal, **45**, (2003), 554–560.
- R. A. Bailey and C. A. Rowley: Valid randomization.
  Proceedings of the Royal Society of London Series A, **410**, (1987), 105–124.
- F. Yates: *Experimental Design: Selected Papers*. Griffin, 1970.

References: Latinization; Unequal block sizes

- B. Harshbarger and L. L. Davis:
  Latinized rectangular lattices. *Biometrics*, **8**, (1952), 73–84.
- E. R. Williams:
  Row and column designs with contiguous replicates.
  Australian Journal of Statistics, **28**, (1986), 154–163.
- J. A. John and E. R. Williams: t-Latinized designs.
  Australian Journal of Statistics, **40**, (1998), 111–118.
- H. D. Patterson and E. R. Williams:
  A new class of resolvable incomplete block designs.
  Biometrika, **63**, (1976), 83–92.

A Positive Note

After I had finished the previous slides, I re-read the paper "Beyond Latin Squares: A Brief Tour of Row–Column Designs" by Hans-Peter Piepho, Emlyn R. Williams and Volker Michel, published in *Agronomy Journal*, **107**, (2015), 2263–2270.

A Positive Note

After I had finished the previous slides, I re-read the paper "Beyond Latin Squares: A Brief Tour of Row–Column Designs" by Hans-Peter Piepho, Emlyn R. Williams and Volker Michel, published in *Agronomy Journal*, **107**, (2015), 2263–2270.

I was delighted to find that it manages to combine the two approaches.

A Positive Note

After I had finished the previous slides, I re-read the paper "Beyond Latin Squares: A Brief Tour of Row–Column Designs" by Hans-Peter Piepho, Emlyn R. Williams and Volker Michel, published in *Agronomy Journal*, **107**, (2015), 2263–2270.

I was delighted to find that it manages to combine the two approaches.

Last week I was at the XIII Working Seminar on Variety Testing at COBORU.

A Positive Note

After I had finished the previous slides, I re-read the paper "Beyond Latin Squares: A Brief Tour of Row–Column Designs" by Hans-Peter Piepho, Emlyn R. Williams and Volker Michel, published in *Agronomy Journal*, **107**, (2015), 2263–2270.

I was delighted to find that it manages to combine the two approaches.

Last week I was at the XIII Working Seminar on Variety Testing at COBORU. In her talk, Karen Wolf mentioned the principle "Analyse as randomized" citing the paper
"Analyze as randomized—why dropping block effects in randomized experiments is a bad idea"
by J. Frey, J. Hartung, J. Ogutu and H.-P. Piepho,
published in *Agronomy Journal*, **116**, (2024), 1371–1381.

A Negative Note

Speakers Jens Hartung and Karen Wolf gave many examples of bad practice in recent agricultural experiments.

A Negative Note

Speakers Jens Hartung and Karen Wolf gave many examples of bad practice in recent agricultural experiments.

1. One series of trials used a different field each year but all trials had exactly the same layout for 15 years.

A Negative Note

Speakers Jens Hartung and Karen Wolf gave many examples of bad practice in recent agricultural experiments.

1. One series of trials used a different field each year but all trials had exactly the same layout for 15 years.

   "We randomized it the first time, so of course it is randomized."

A Negative Note

Speakers Jens Hartung and Karen Wolf gave many examples of bad practice in recent agricultural experiments.

1. One series of trials used a different field each year but all trials had exactly the same layout for 15 years.

   "We randomized it the first time, so of course it is randomized."

2. "We tested variety A in 7 trials in Region 1;
   we tested variety B in 7 trials in Region 2."

A Negative Note

Speakers Jens Hartung and Karen Wolf gave many examples of bad practice in recent agricultural experiments.

1. One series of trials used a different field each year but all trials had exactly the same layout for 15 years.

   "We randomized it the first time, so of course it is randomized."

2. "We tested variety A in 7 trials in Region 1; we tested variety B in 7 trials in Region 2."

   There is no way of distinguishing the difference between the 2 varieties from the difference between the 2 regions.

A Negative Note

Speakers Jens Hartung and Karen Wolf gave many examples of bad practice in recent agricultural experiments.

1. One series of trials used a different field each year but all trials had exactly the same layout for 15 years.

   "We randomized it the first time, so of course it is randomized."

2. "We tested variety A in 7 trials in Region 1; we tested variety B in 7 trials in Region 2."

   There is no way of distinguishing the difference between the 2 varieties from the difference between the 2 regions. This is a common mistake, known as pseudo-replication.

A Negative Note

Speakers Jens Hartung and Karen Wolf gave many examples of bad practice in recent agricultural experiments.

1. One series of trials used a different field each year but all trials had exactly the same layout for 15 years.

   “We randomized it the first time, so of course it is randomized.”

2. “We tested variety A in 7 trials in Region 1; we tested variety B in 7 trials in Region 2.”

   There is no way of distinguishing the difference between the 2 varieties from the difference between the 2 regions. This is a common mistake, known as pseudo-replication.

3. “What are the blocks in this experiment?”

A Negative Note

Speakers Jens Hartung and Karen Wolf gave many examples of bad practice in recent agricultural experiments.

1. One series of trials used a different field each year but all trials had exactly the same layout for 15 years.

   "We randomized it the first time, so of course it is randomized."

2. "We tested variety A in 7 trials in Region 1; we tested variety B in 7 trials in Region 2."

   There is no way of distinguishing the difference between the 2 varieties from the difference between the 2 regions. This is a common mistake, known as pseudo-replication.

3. "What are the blocks in this experiment?"

   "We do not know: we just copied the design and layout from someone else's experiment that they did last year. We have no idea what they mean by 'blocks'."

Looking Back and Forward

When I worked at Rothamsted in the 1980s,
no experiment was allowed to be started until its design had been approved by a statistician

Looking Back and Forward

When I worked at Rothamsted in the 1980s,
no experiment was allowed to be started until its design had been approved by a statistician
(who also ran a dummy statistical analysis on dummy data to make sure that there was no hidden confounding).

Looking Back and Forward

When I worked at Rothamsted in the 1980s,
no experiment was allowed to be started until its design had been approved by a statistician
(who also ran a dummy statistical analysis on dummy data to make sure that there was no hidden confounding).

Nowadays, statisticians are often not consulted until the experiment has been done and the data collected.

Looking Back and Forward

When I worked at Rothamsted in the 1980s,
no experiment was allowed to be started until its design had been approved by a statistician
(who also ran a dummy statistical analysis on dummy data to make sure that there was no hidden confounding).

Nowadays, statisticians are often not consulted until the experiment has been done and the data collected.
Then it is too late to correct many of these basic mistakes.

Looking Back and Forward

When I worked at Rothamsted in the 1980s,
no experiment was allowed to be started until its design had been approved by a statistician
(who also ran a dummy statistical analysis on dummy data to make sure that there was no hidden confounding).

Nowadays, statisticians are often not consulted until the experiment has been done and the data collected.
Then it is too late to correct many of these basic mistakes.

I do hope that some of the people in this room can explain to experimenters how to improve their approach.