Multiple randomizations
(Brien and Bailey, 2006, *JRSS B*, **68**, 571-609; 2008?; 2009?)

Chris Brien

School of Mathematics and Statistics

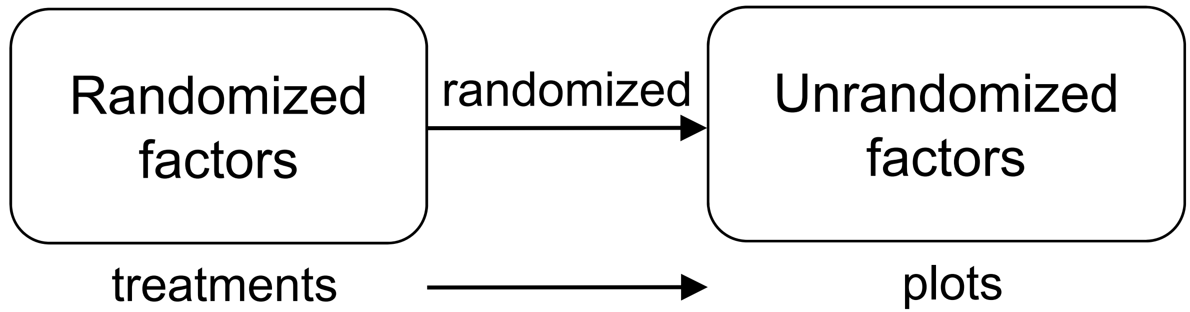
University of South Australia

chris.brien@unisa.edu.au

Outline

1. Some history and a simple example
2. Overview of types of multiple randomizations
3. Chains of randomizations
4. Two sets randomized to one
5. One set randomized to two
6. Designing multitiered experiments
7. Pseudofactors
8. Summary

A randomization: its elements



Sets of objects	Γ (treatments)	Ω (plots)
Tiers: sets of factors	$\mathcal{F}_\Gamma$ (Randomized)	$\mathcal{F}_\Omega$ (Unrandomized)
Vector spaces	$V_\Gamma \leq V_\Omega$	$V_\Omega = R^n$
Orthogonal decompositions (structures)	$\mathcal{R}$, a set of **R**s	$\mathcal{U}$, a set of **U**s

Start with $\mathcal{U}$ and further decompose according to $\mathcal{R}$.

The decomposition

- Consider structure-balanced experiments, those for which:
  - $\mathbf{RUR} = \lambda_{\mathbf{UR}}\mathbf{R}$ for all $\mathbf{R}\in\mathcal{R}$, $\mathbf{U}\in\mathcal{U}$
    ($0\leq\lambda_{\mathbf{UR}}\leq 1$ is called a canonical efficiency factor and is the nonzero eigenvalue of both **RUR** and **URU**.)
  - $\mathbf{R}_1\mathbf{U}\mathbf{R}_2 = 0$ for all $\mathbf{U}\in\mathcal{U}$ and $\mathbf{R}_1 \neq \mathbf{R}_2$
- The complete decomposition specified by $\mathcal{U}$ and its further decomposition according to $\mathcal{R}$ is given by

$$\mathcal{U} \rhd \mathcal{R} = \left\{\mathbf{U} \rhd \mathbf{R} : \mathbf{U} \in \mathcal{U}, \mathbf{R} \in \mathcal{R}, \lambda_{\mathbf{UR}} \neq 0\right\} \cup \left\{\mathbf{U} \vdash \mathcal{R} : \mathbf{U} \in \mathcal{U}\right\}$$

where

$$\mathbf{U} \rhd \mathbf{R} = \lambda_{\mathbf{UR}}^{-1}\mathbf{URU}$$ **U** pertaining to **R**

$$\mathbf{U} \vdash \mathcal{R} = \mathbf{U} - \sum\nolimits_{\mathbf{R}\in\mathcal{R}} \lambda_{\mathbf{UR}}^{-1}\mathbf{URU}$$ **U** orthogonal to all **R**

- Not necessarily confined to structure-balanced experiments, but $\mathbf{U}\rhd\mathbf{R}$ not so simple.

1 Some history

1955: McIntyre introduced two-phase experiments using a TMV experiment with tester plants

1958: Cox recognized the need for multiple randomizations and gave some examples (Brien & Bailey, 2006, Ex. 4 Cotton fibres)

1959: Curnow corrected McIntyre's analysis

1983: Brien published a sensory experiment and introduced the term multitiered.

1988: Wood, Williams and Speed looked at nonorthogonal variance structure in two-phase experiments

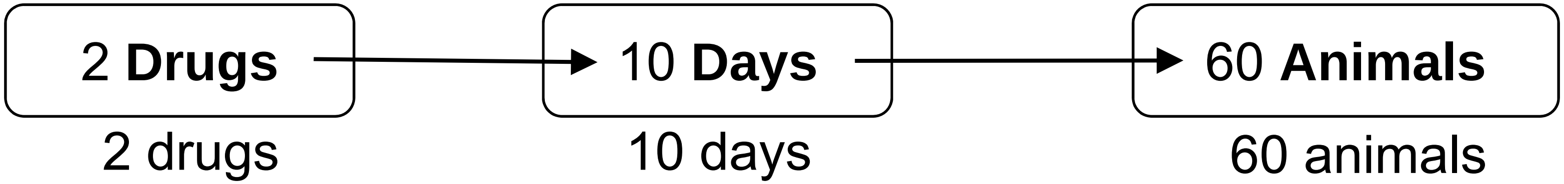
1999: Brien and Payne give ANOVA algorithm

2003: Cullis et al discuss multiphase nature of barley malting trials

2003: Kerr mentions that microarrays experiments can be two-phase

2006: Brien and Bailey paper on multiple randomizations

White's (1975) simple animal experiment



animals tier		days tier		drugs tier	
source	df	source	df	source	df
Mean	1	Mean	1	Mean	1
Animals	59	Days	9	Drugs	1
				Residual	8
		Residual	50		

- Table shows that Drugs differences confounded with Days differences and Animals Differences.
- The Residual has 8 df and gives the variability of Animals plus Days.
- Fundamentally different to:
  - grouping Animals based on Days and
  - randomizing Drugs to Days.

2. Overview of types of multiple randomizations

- Two randomizations with three sets of objects: Ω, Υ, Γ
- There are tiers of factors indexing each: $\mathcal{F}_{\Omega}$, $\mathcal{F}_{\Upsilon}$, $\mathcal{F}_{\Gamma}$.

Chain: one arrow follows another

Composed

$$\Gamma \longrightarrow \Upsilon \longrightarrow \Omega$$

Randomized-inclusive

$$[\Gamma \longrightarrow \Upsilon] \longrightarrow \Omega$$

Same start: one to two:

Double

$$\Upsilon \longleftarrow \Gamma \longrightarrow \Omega$$

Same end: two to one

Independent

$$\Upsilon \longrightarrow \Omega \longleftarrow \Gamma$$

Coincident

$$\Upsilon \longrightarrow \Omega \longleftarrow \Gamma$$

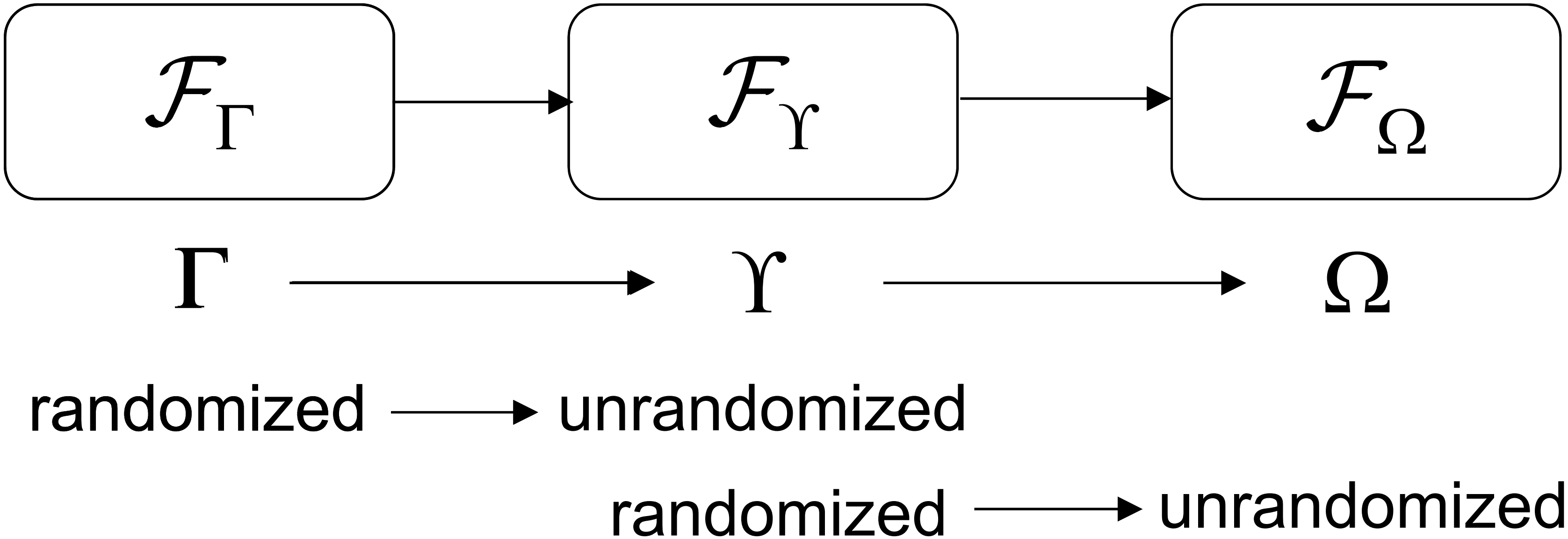
Unrandomized-inclusive

$$[\Upsilon \longrightarrow \Omega] \longleftarrow \Gamma$$

- Order matters only for the inclusive randomizations

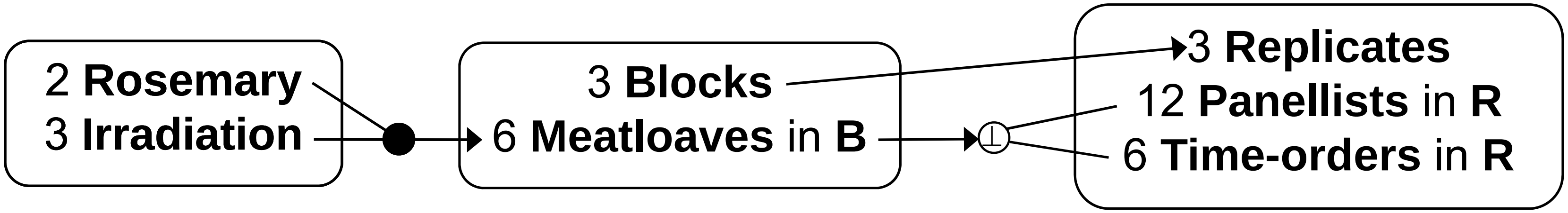
3. Chain of randomizations

a) Composed randomizations: Order does not matter



A two-phase sensory experiment

(Brien & Bailey, 2006, Fig. 7; T. B. Bailey, 2003)



6 treatments

18 meatloaves

216 tastings

Complete-block design

Two 6 x 6 Latin squares

tastings tier		meatloaves tier		treatments tier	
source	df	source	df	source	df
Mean	1	Mean	1	Mean	1
Replicates	2	Blocks	2		
Panellists[Rep]	33				
Time-orders[Rep]	15				
P#T[Rep]	165	Meatloaves[B]	15	Rosemary	1
				Irradiation	2
				R#I	2
				Residual	10
		Residual	150		

Composed randomizations: elements

	treatments $\longrightarrow$	1^{st} units $\longrightarrow$	2^{nd} units
Sets of objects	Γ	Υ	Ω
Tiers	$\mathcal{F}_{\Gamma}$	$\mathcal{F}_{\Upsilon}$	$\mathcal{F}_{\Omega}$
Vector spaces	$V_{\Gamma} \leq V_{\Upsilon} \leq V_{\Omega}$	$V_{\Upsilon} \leq V_{\Omega}$	$V_{\Omega} = R^{n}$
Orthogonal decompositions (structures)	$\mathcal{T}$, a set of **T**s	$\mathcal{R}$, a set of **R**s	$\mathcal{U}$, a set of **U**s

- CJB approach (analysis): $(\mathcal{U} \rhd \mathcal{R}) \rhd \mathcal{T}$
  - start with $\mathcal{U}$, decompose according to $\mathcal{R}$, and then further decompose according to $\mathcal{T}$.
- RAB approach (design): $\mathcal{U} \rhd (\mathcal{R} \rhd \mathcal{T})$
  - start with $\mathcal{R}$, decompose according to $\mathcal{T}$, and then decompose $\mathcal{U}$ according to $\mathcal{R}$ and $\mathcal{T}$.

Theorem for composed randomizations

(Brien & Bailey, 2008)

Treatments →	1st units →	2nd units
$\mathcal{T}$	$\mathcal{R}$	$\mathcal{U}$

- If $\mathcal{T}$ is structure balanced in relation to $\mathcal{R}$ (with efficiency matrix $\Lambda_{\mathcal{RT}}$); and $\mathcal{R}$ is structure balanced in relation to $\mathcal{U}$ (with efficiency matrix $\Lambda_{\mathcal{UR}}$), then
  - $\mathcal{T}$ is structure balanced in relation to $\mathcal{U}$ (with efficiency matrix $\Lambda_{\mathcal{UT}} = \Lambda_{\mathcal{UR}} \Lambda_{\mathcal{RT}}$);
  - $\mathcal{T}$ is structure balanced in relation to $\mathcal{U} \rhd \mathcal{R}$ (with $\Lambda\ldots$);
  - $\mathcal{R} \rhd \mathcal{T}$ is structure balanced in relation to $\mathcal{U}$ (with $\Lambda\ldots$);
  - $(\mathcal{U} \rhd \mathcal{R}) \rhd \mathcal{T} = \mathcal{U} \rhd (\mathcal{R} \rhd \mathcal{T})$.
- Implication for two-phase experiments is:
  - treat each phase's design independently, using a structure-balanced design for each phase
  - complete decomposition obtained by getting the two-tiered decomposition $\mathcal{U} \rhd \mathcal{R}$ first, and then refining this for $\mathcal{T}$ — Decomposition table mimics this.

Idempotents for a two-phase sensory experiment

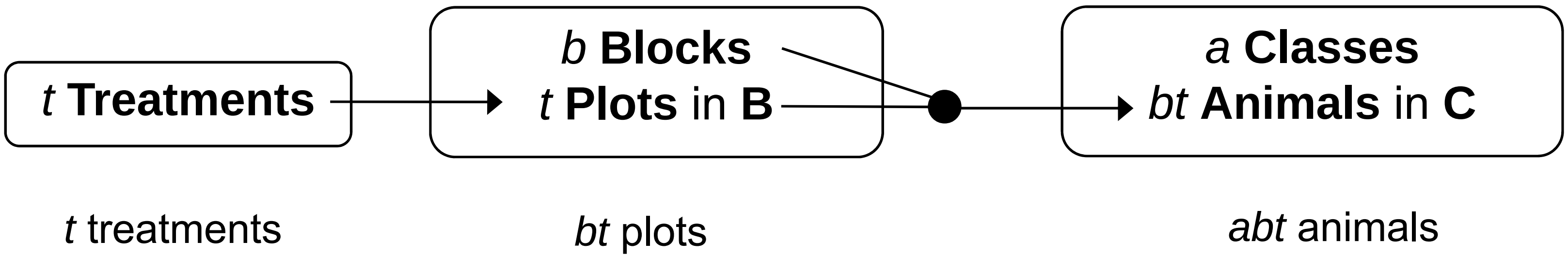
tastings tier		meatloaves tier		treatments tier	
source	$\mathcal{U}$	source	$\mathcal{U}\rhd\mathcal{R}$	source	$\mathcal{U}\rhd\mathcal{R}\rhd\mathcal{T}$
Mean	$\mathbf{U}_{Mean}$	Mean	$\mathbf{U}_{Mean}\rhd\mathbf{R}_{Mean}$	Mean	$\mathbf{U}_{Mean}\rhd\mathbf{R}_{Mean}\rhd\mathbf{T}_{Mean}$
Rep	$\mathbf{U}_{Rep}$	B	$\mathbf{U}_{Rep}\rhd\mathbf{R}_{B} = \mathbf{R}_{B}$		
P[Rep]	$\mathbf{U}_{P[Rep]}$				
T[Rep]	$\mathbf{U}_{T[Rep]}$				
P#T[Rep]	$\mathbf{U}_{P\#T[Rep]}$	M[B]	$\mathbf{U}_{P\#T[Rep]}\rhd\mathbf{R}_{M[B]} = \mathbf{R}_{M[B]}$	R	$\mathbf{U}_{P\#T[Rep]}\rhd\mathbf{R}_{M[B]}\rhd\mathbf{T}_{R} = \mathbf{T}_{R}$
				I	$\mathbf{U}_{P\#T[Rep]}\rhd\mathbf{R}_{M[B]}\rhd\mathbf{T}_{I} = \mathbf{T}_{I}$
				R#I	$\mathbf{U}_{P\#T[Rep]}\rhd\mathbf{R}_{M[B]}\rhd\mathbf{T}_{R\#I} = \mathbf{T}_{R\#I}$
				Residual	$\mathbf{U}_{P\#T[Rep]}\rhd\mathbf{R}_{M[B]}\vdash\mathcal{T}$
		Residual	$\mathbf{U}_{P\#T[Rep]}\vdash\mathbf{R}_{M[B]}$		

$\mathcal{R} = \{\mathbf{R}_{Mean}, \mathbf{R}_{B}, \mathbf{R}_{M[B]}\}$ $\mathcal{T} = \{\mathbf{T}_{Mean}, \mathbf{T}_{R}, \mathbf{T}_{I}, \mathbf{T}_{R\#I}\}$

- P = Panellists; T = Time-orders;
- B = Blocks; M = Meatloaves
- R = Rosemary, I = Irradiation

A continuous grazing experiment

(Brien & Demétrio, 1998; Brien & Bailey, 2006, ex.3)



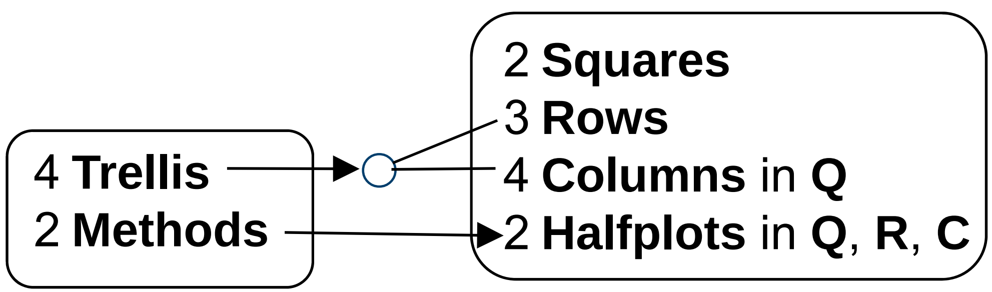
- A single-phase experiment with two randomizations
- Two EUs: Plots in B and Animals in C
- What is the decomposition table?

animals tier		plots tier		treatments tier	
source	df	source	df	source	df
Mean	1	Mean	1	Mean	1
Classes	$a-1$				
Animals[C]	$a(bt-1)$	Blocks	$b-1$		
		Plots[B]	$b(t-1)$	Treatments	$t-1$
				Residual	$(b-1)(t-1)$
		Residual	$(a-1)(bt-1)$		

A nonorthogonal, structure-balanced sensory experiment (Brien & Payne, 1999, Appl. Stats)

- Field phase — a split-plot viticultural experiment
  - For main plots, 4 trellises randomized using 2 adjacent Youden squares:

Squares		1				2			
Columns		1	2	3	4	1	2	3	4
	Rows								
	1	4	1	2	3	2	1	4	3
	2	1	2	3	4	3	2	1	4
	3	2	3	4	1	4	3	2	1



(Q = Squares)

8 treatments

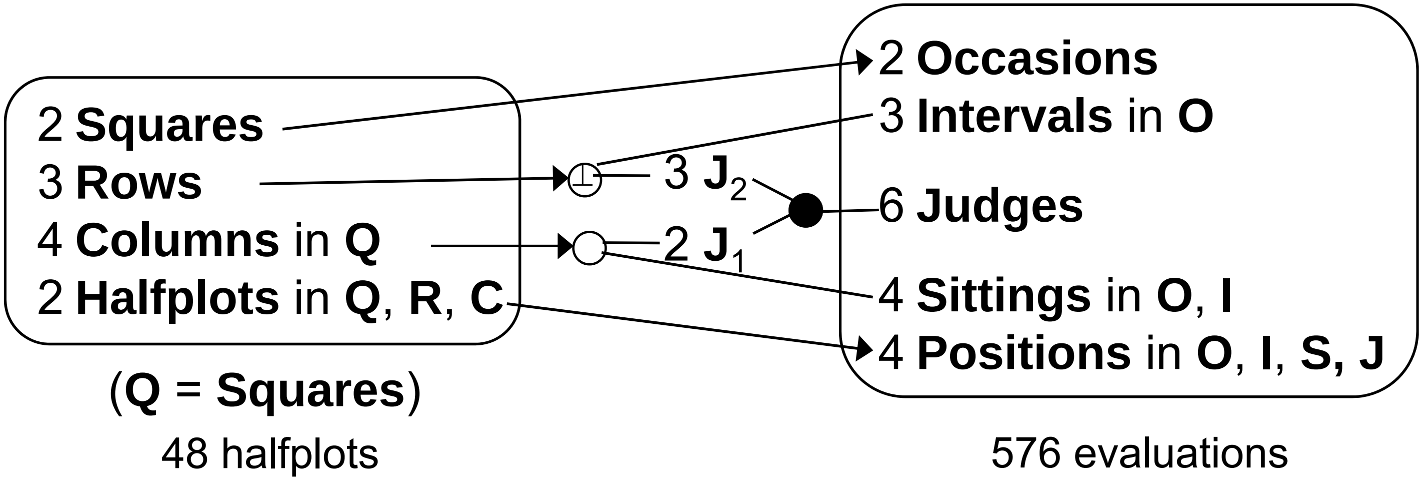
48 halfplots

  - For halfplots, 2 pruning methods randomized to them.
  - Produce of each halfplot was made into wine ⇒ 48 wines

The structure-balanced sensory experiment (cont'd)

- Evaluation phase
  - 48 wines were scored by 6 judges
  - All took part in 24 sittings divided into:

    4 Sittings within 3 Intervals within 2 Occasions
  - 2 Squares randomized to 2 Occasions — same for all judges
  - 2 Half-plots from a main plot randomized to wine glasses in 4 bench positions at each sitting so each wine had 2 replicate glasses
  - Main plot (row-column) randomization uses 3 × 3 Latin squares and 3 × 2 × 4 extended Youden squares



Assignment of rows to Judges (within sets) × Sittings

	Intervals	1				2				3			
	Sittings	1	2	3	4	1	2	3	4	1	2	3	4
	Judges												
Occasion 1	1	1	1	1	1	3	3	3	3	2	2	2	2
	2	2	2	2	2	1	1	1	1	3	3	3	3
	3	3	3	3	3	2	2	2	2	1	1	1	1
	4	3	3	3	3	2	2	2	2	1	1	1	1
	5	1	1	1	1	3	3	3	3	2	2	2	2
	6	2	2	2	2	1	1	1	1	3	3	3	3
Occasion 2	1	2	2	2	2	3	3	3	3	1	1	1	1
	2	1	1	1	1	2	2	2	2	3	3	3	3
	3	3	3	3	3	1	1	1	1	2	2	2	2
	4	3	3	3	3	1	1	1	1	2	2	2	2
	5	2	2	2	2	3	3	3	3	1	1	1	1
	6	1	1	1	1	2	2	2	2	3	3	3	3

Two 3×3 Latin squares on each occasion:

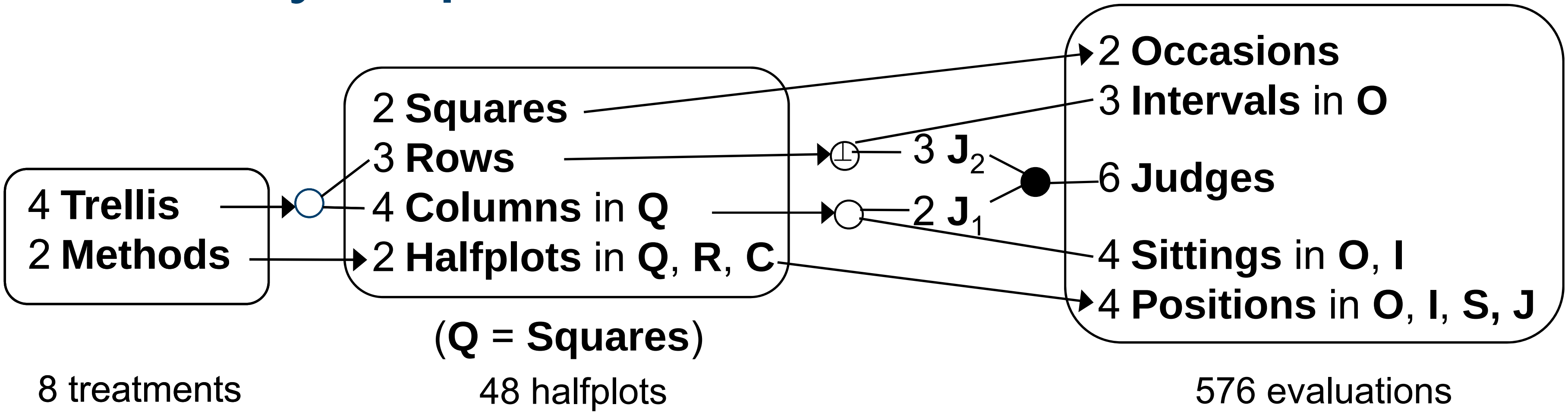
⇒ one for judges 1–3, other for judges 4 –6

Assignment of Columns to Judges (between sets) × Intervals

	Intervals	1				2				3			
	Sittings	1	2	3	4	1	2	3	4	1	2	3	4
	Judges												
Occasion 1	1	3	2	1	4	1	4	2	3	2	3	4	1
	2	3	2	1	4	1	4	2	3	2	3	4	1
	3	**3**	**2**	**1**	**4**	**1**	**4**	**2**	**3**	**2**	**3**	**4**	**1**
	4	**1**	**4**	**3**	**2**	**2**	**3**	**1**	**4**	**3**	**2**	**1**	**4**
	5	1	4	3	2	2	3	1	4	3	2	1	4
	6	1	4	3	2	2	3	1	4	3	2	1	4
Occasion 2	1	4	1	2	3	1	3	2	4	1	3	2	4
	2	4	1	2	3	1	3	2	4	1	3	2	4
	3	**4**	**1**	**2**	**3**	**1**	**3**	**2**	**4**	**1**	**3**	**2**	**4**
	4	**3**	**2**	**1**	**4**	**3**	**1**	**4**	**2**	**4**	**2**	**3**	**1**
	5	3	2	1	4	3	1	4	2	4	2	3	1
	6	3	2	1	4	3	1	4	2	4	2	3	1

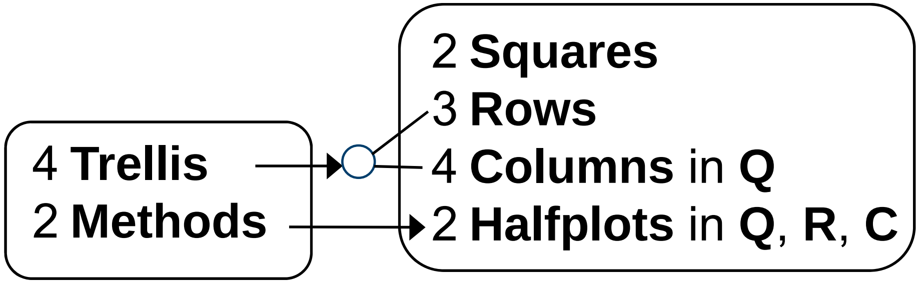
- Uses a 3 × 2 × 4 extended Youden design on each occasion

Composed randomizations in the sensory experiment



- A chain of two randomizations:
  - treatments to wines and wines to positions.
- Three-tiered: (i) treatments: randomized field (& eval indirectly);
  (ii) wines: unrandomized field, randomized eval;
  (iii) positions: unrandomized eval.
- Model formulae:
  - ((Occasions / Intervals / Sittings) * Judges) / Positions
  - (Rows * (Squares / Columns)) / Halfplots
  - Trellis * Methods

Field-phase decomposition table



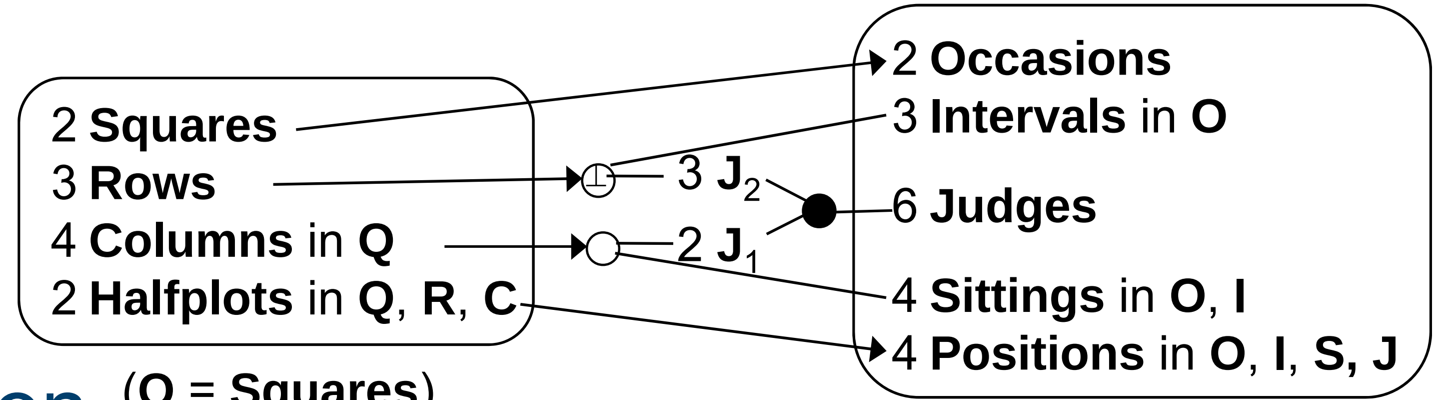
(**Q** = **Squares**)

8 treatments

48 halfplots

wines tier		treatments tier		
source	df	eff	source	df
Squares	1			
Rows	2			
Q#R	2			
Columns[Q]	6	1/9	Trellis	3
			Residual	3
R#C[Q]	12	8/9	Trellis	3
			Residual	9
Halfplots[R∧C∧Q]	24		Method	1
			T#M	3
			Residual	20

Evaluation-phase decomposition table



(**Q** = **Squares**)

48 halfplots positions tier			576 evaluations wines tier	
source	df	eff	source	df
Occasions	1		Squares	1
Judges	5			
O#J	5			
Intervals[O]	4			
I#J[O]	20		Rows	2
			Q#R	2
			Residual	16
Sittings[O∧I]	18	1/3	Columns[Q]	6
			Residual	12
S#J[O∧I]	90	2/3	Columns[Q]	6
			R#C[Q]	12
			Residual	72
Positions[O∧I∧S∧J]	432		Halfplots[R∧C∧Q]	24
			Residual	408

Combined decomposition table

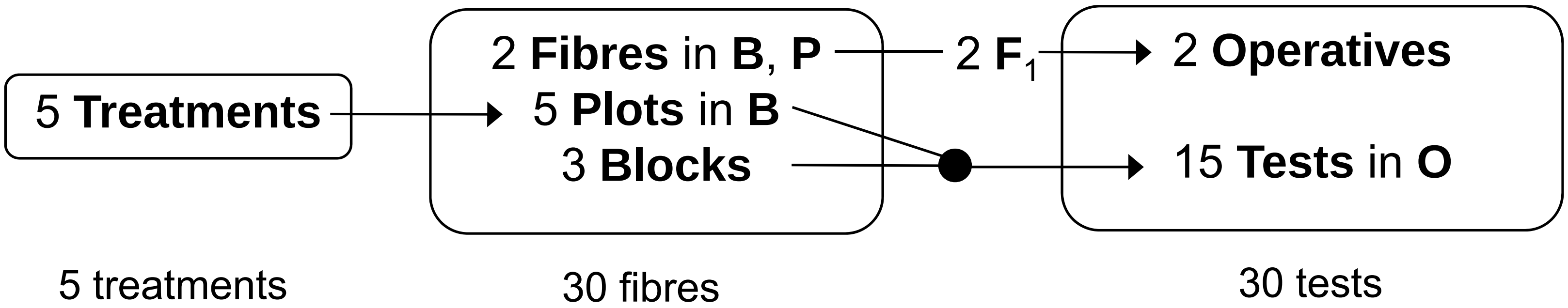
positions tier		wines tier			treatments tier		
source	df	eff	source	df	eff	source	df
Occasions	1		Squares	1			
Judges	5						
O#J	5						
Intervals[O]	4						
I#J[O]	20		Rows	2			
			Q#R	2			
			Residual	16			
Sittings[O∧I]	18	1/3	Columns[Q]	6	1/27	Trellis	3
						Residual	3
			Residual	12			
S#J[O∧I]	90	2/3	Columns[Q]	6	2/27	Trellis	3
						Residual	3
			R#C[Q]	12	8/9	Trellis	3
						Residual	9
			Residual	72			
Positions[O∧I∧S∧J]	432		Halfplots[R∧C∧Q]	24		Method	1
						T#M	3
						Residual	20
			Residual	408			

Composed randomizations in the sensory experiment

- Efficiency factors for Trellis are product of those from the designs for the two phases.
- Results in most Trellis information randomized to Judges#Sittings [O∧I] so major effects arising from Occasions, Intervals, Judges and Sittings eliminated.
- Residual df for Trellis determined by field-phase design.
- Methods only affected by variability amongst glasses in 4 positions evaluated together.

Cotton fibres (D. R. Cox, 1958; Brien and Bailey, 2006, Ex. 4)

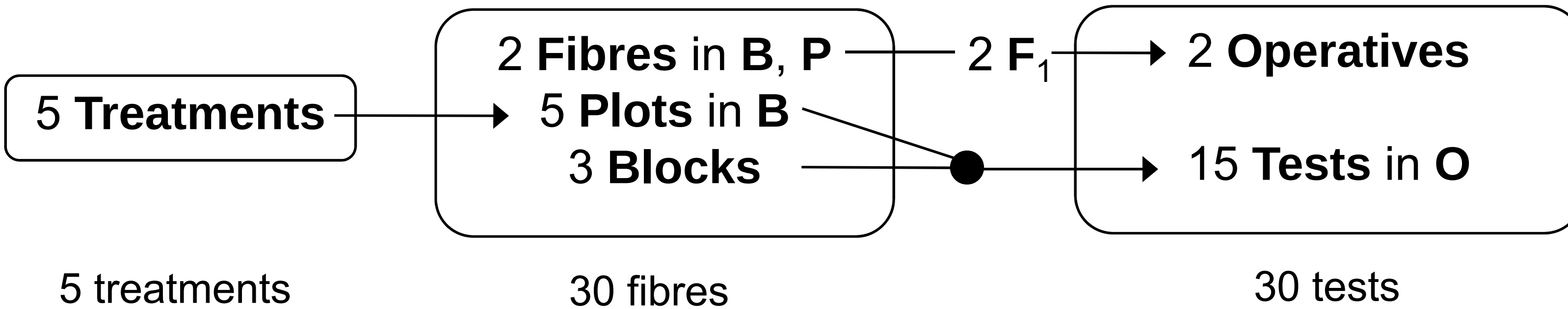
- Field and testing phases



- F_1 is a **pseudofactor**
  - two levels, each indexing half the fibres
  - levels randomized independently in each of 15 plots
  - no inherent meaning
  - shown outside the panel
- Randomization is not **consonant** because:
  - Fibres are nested in Blocks∧Plots: 30 fibres
  - Operatives are not nested in Tests: 2 Operatives

Cotton fibres (continued)

- Field and testing phases



tests tier		fibres tier		treatments tier	
source	df	source	Df	source	df
Operatives	1	F_1	1		
Tests[O]	28	Blocks	2		
		Plots[B]	12	Treatments	4
				Residual	8
		Fibres[B∧P]⊢F_1	14		

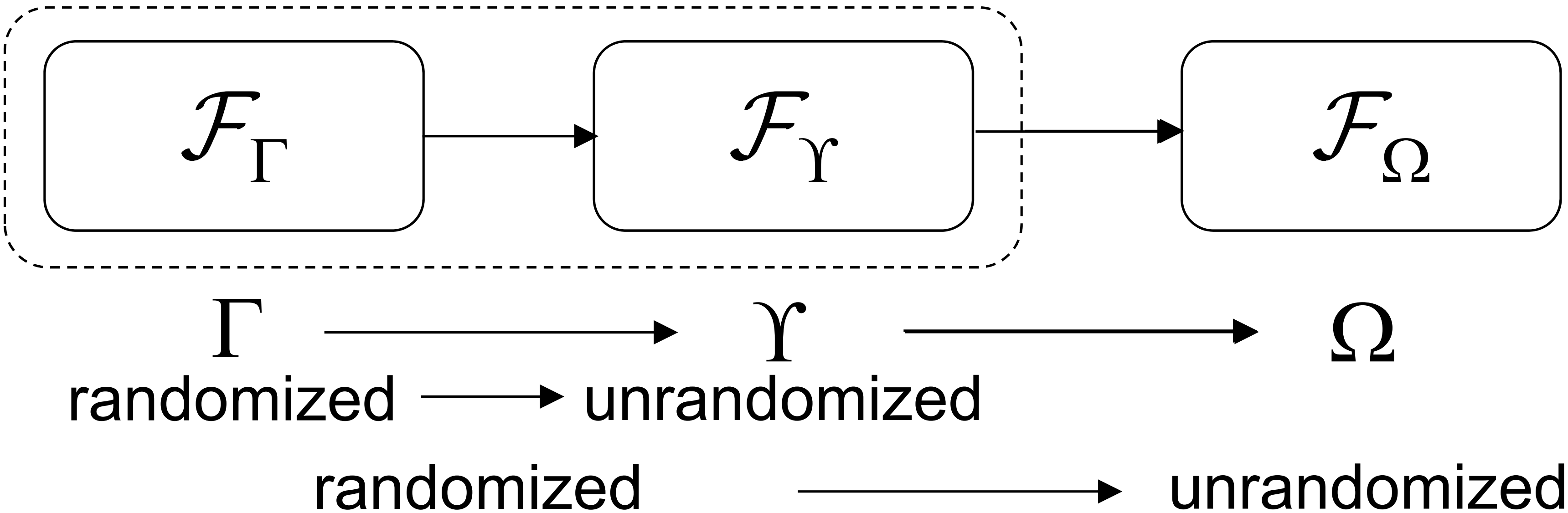
A simple example

- Field phase
  - RCBD in which 3 levels of N x 2 levels of P assigned to 6 plots in 5 blocks
- Laboratory phase
  - RCBD for 6 analyses in 5 occasions in which blocks assigned to occasions and plots to analyses.

- What are the sets of objects?
- What is the randomization diagram?
- What are the tiers?
- What are the model formulae and sources?
- What is the decomposition table, including the idempotents?

3. Chain of randomizations

b) Randomized-inclusive randomization: Order does matter



- Two tiers from 1st randomization are randomized jointly and form a pseudotier
  - ➢ Factors have same status in THIS randomization only.
- Need information from **both** tiers from 1st when doing 2nd randomization.
  - ➢ Typically information encoded in pseudofactors.

Decomposition for randomized-inclusive randomizations

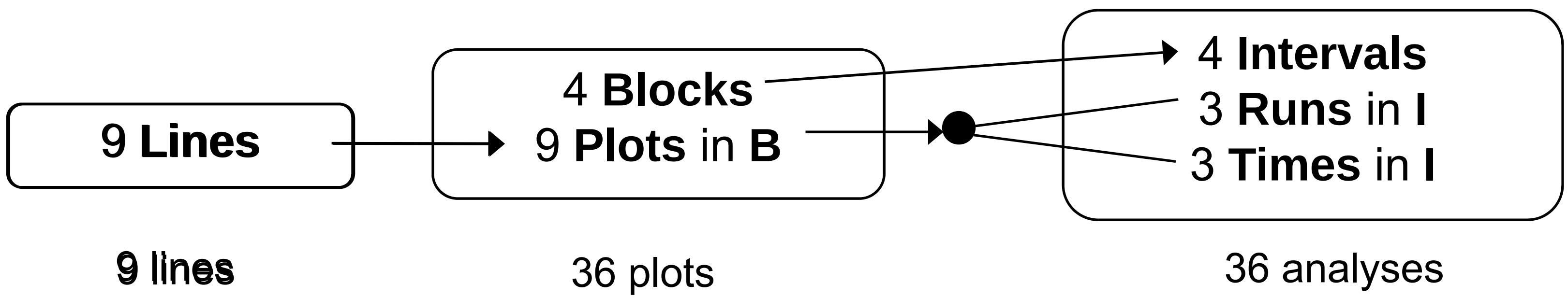
Treatments →	1st units →	2nd units
Γ	Υ	Ω
$\mathcal{F}_\Gamma$	$\mathcal{F}_\Upsilon$	$\mathcal{F}_\Omega$
$\mathcal{T}$	$\mathcal{R}$	$\mathcal{U}$

- Similar set up to composed randomizations
- Generally, the first randomization uses a design so that $\mathcal{T}$ is structure balanced in relation to $\mathcal{R}$ (with efficiency matrix $\Lambda_{\mathcal{RT}}$);
- However, not possible to obtain a design for which $\mathcal{R}$, derived from $\mathcal{F}_\Upsilon$, is structure balanced in relation to $\mathcal{U}$ without the introduction of pseudofactors.
- The pseudofactors introduce new elements that refine the original elements of $\mathcal{R}$ so that $\mathcal{R}$, derived from $\mathcal{F}_\Upsilon$, is structure balanced in relation to $\mathcal{U}$ (with efficiency matrix $\Lambda_{\mathcal{UR}}$).
- Then the results for composed randomizations apply, e.g.
  - $\mathcal{T}$ is structure balanced in relation to $\mathcal{U}$ (with efficiency matrix $\Lambda_{\mathcal{UT}} = \Lambda_{\mathcal{UR}}\ \Lambda_{\mathcal{RT}}$);
  - $(\mathcal{U} \triangleright \mathcal{R}) \triangleright \mathcal{T} = \mathcal{U} \triangleright (\mathcal{R} \triangleright \mathcal{T})$.

A Two-Phase Wheat Variety Trial

(like Brien & Bailey, 2006, Example 9)

- Involves two randomizations:
  - *Field phase:* 9 lines to 4 blocks of 9 plots using an RCBD.
  - *Laboratory phase:* sample from each of 36 plots analysed in gas chromatograph that processes 3 samples per run.



- Composed randomizations but Lines likely nonorthogonal.
- Need r-inclusive randomization because
  i. **Lines** randomized to **Plots** in **B**, and in turn,
  ii. **Plots** in **B** is randomized to a) **Runs** in **I** and b) **Times** in **I**, such that,
  iii. Plots within Blocks not balanced w.r.t. any one of Runs, Times or Runs-Times within Intervals (λ = 1 only for those contrasts confounded with a source; 0 otherwise).

A 3×3 balanced lattice square

Square	I			II		
Cols	1	2	3	1	2	3
Rows						
1	1	2	3	1	6	8
2	4	5	6	9	2	4
3	7	8	9	5	7	3

Square	III			IV		
Rows						
1	1	4	7	1	9	5
2	2	5	8	6	2	7
3	3	6	9	8	4	3

Rows and columns from first two squares interchanged in last two squares

- Use to work out Lines and hence Plots in B to randomize to Runs-Times.
  - Think of numbers in a row as specifying that the plots with these Lines are to be placed in the same Run

Lab randomization of 2 (of 4) Blocks using a balanced lattice square

Blocks	Plots	P_1	P_2	L_1	L_2	L_3	L_4	Lines
1	5	1	1	1	1	1	1	1
	6	1	2	1	2	2	2	2
	9	1	3	1	3	3	3	3
	8	2	1	2	1	2	3	4
	7	2	2	2	2	3	1	5
	4	2	3	2	3	1	2	6
	2	3	1	3	1	3	2	7
	3	3	2	3	2	1	3	8
	1	3	3	3	3	2	1	9
2	3	1	1	1	1	1	1	1
	7	1	2	2	3	1	2	6
	8	1	3	3	2	1	3	8
	6	2	1	3	3	2	1	9
	1	2	2	1	2	2	2	2
	2	2	3	2	1	2	3	4
	4	3	1	2	2	3	1	5
	9	3	2	3	1	3	2	7
	5	3	3	1	3	3	3	3

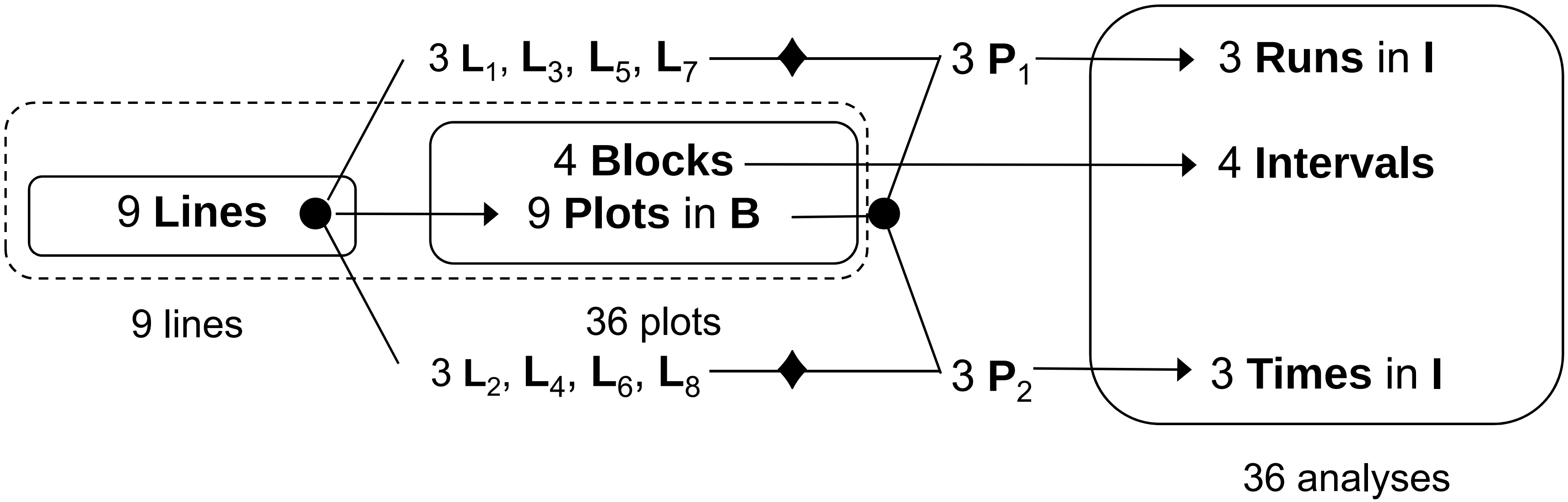
Randomize:

$\mathbf{P_1} \times \mathbf{P_2}$ in **B**,

along with $\mathbf{L_1}, \ldots, \mathbf{L_4}$,

to **Runs**×**Times** in **I**

R-inclusive randomizations in a balanced, two-phase wheat experiment (cont'd)



- Involves 2 randomizations: lines to plots; lines-plots to analyses.
- **Randomized** factors for 2nd randomization **inclusive** of both randomized and unrandomized factors from 1st randomization
  - ➢ randomized-inclusive randomizations

Decomposition table for wheat experiment

analyses tier

source	df
Intervals	3
Runs[I]	8
Times[I]	8
T#I[I]	16

plots tier

source	df
Blocks	3
Plots[B]	32

This plots decomposition is not structure-balanced with respect to the analyses decomposition.

analyses tier		plots tier	
source	df	source	df
Intervals	3	Blocks	3
Runs[I]	8	P_1[B]	8
Times[I]	8	P_2[B]	8
T#I[I]	16	Plots[B]$_{\vdash}$	16

However, this plots decomposition is. The pseudofactors have identified 3 subspaces that give a balanced decomposition.

Decomposition table for wheat experiment

analyses tier		plots tier		lines tier		
source	df	source	df	eff	source	df
Intervals	3	Blocks	3			
Runs[I]	8	P_1[B]	8	0.25	Lines	8
Times[I]	8	P_2[B]	8	0.25	Lines	8
T#I[I]	16	Plots[B]	16	0.5	Lines	8
					Residual	8

- The use of pseudofactors enables one to keep track of the plots.
- Usually ignore and then uneasy about whether something has been missed

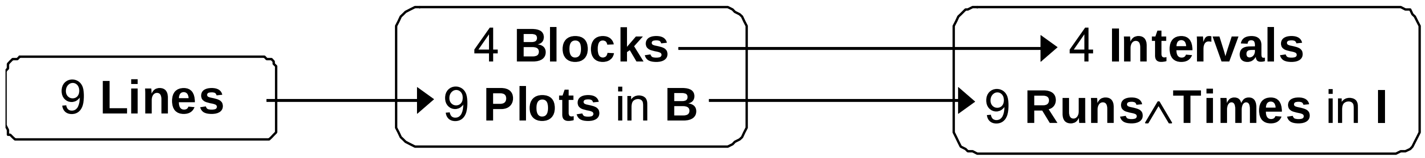
Skeleton ANOVA table for wheat experiment

analyses tier		plots tier		lines tier		
source	df	source	df	eff	source	df
Intervals	3	Blocks	3			
Runs[I]	8	Plots[B]	8	0.25	Lines	8
Times[I]	8	Plots[B]	8	0.25	Lines	8
T#I[I]	16	Plots[B]	16	0.5	Lines	8
					Residual	8

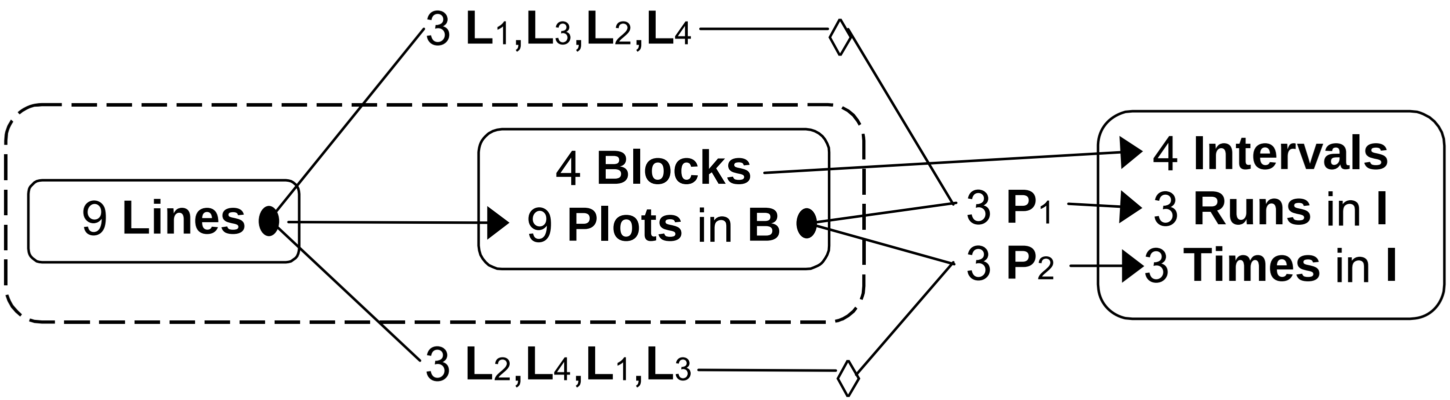
- Final relabelling of sources to reflect their origin.

Key difference between composed and r-inclusive randomizations

Equal vs. unequal efficiency for field sources

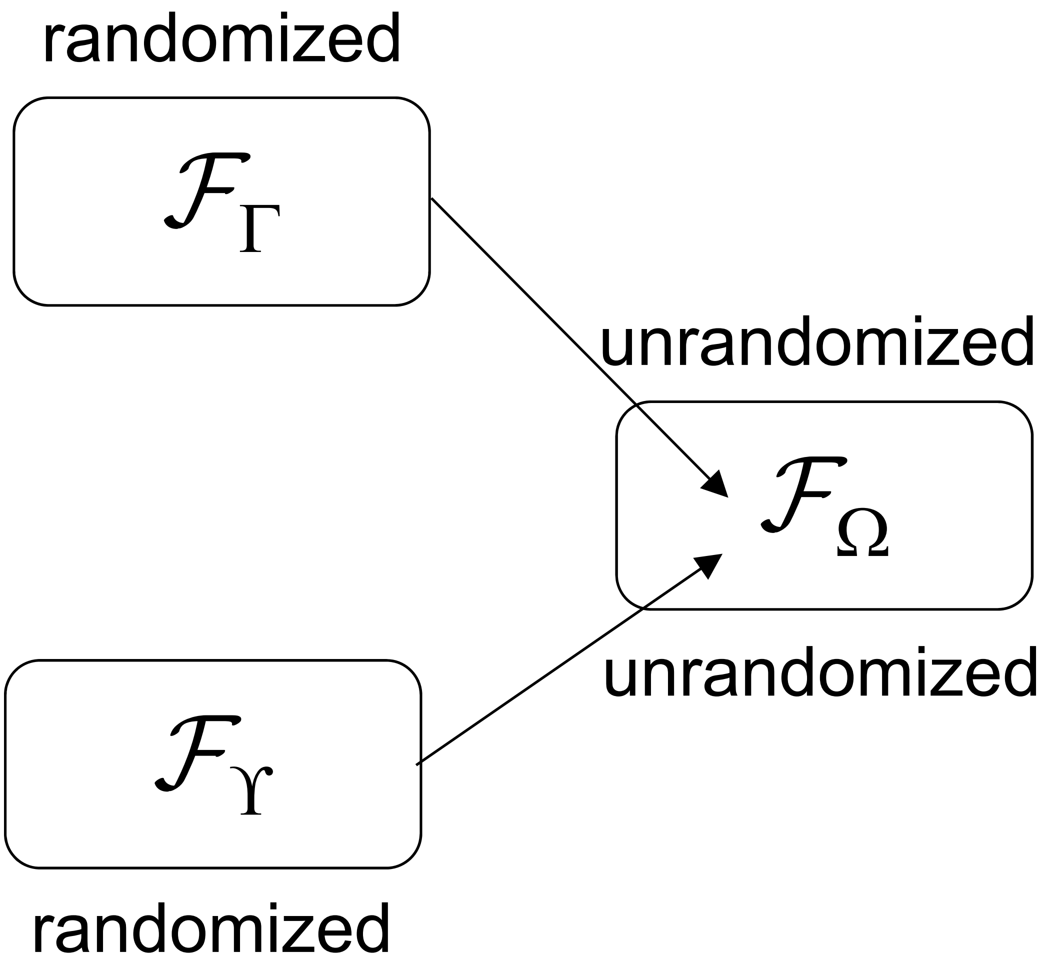


- All 32 df for Plots[Blocks] confounded with R∧T[I]; one *e* = 1.
- No knowledge of Lines randomization needed for 2nd randomization.



- 32 df for Plots[Blocks] divided into subspaces of 8, 8 and 16 df for confounding with R[I], T[I] and R#T[I].
  - e_i for 3 P[B] subspaces with R#T[I] are 0, 0, 1
- Lines randomization determines Plots pseudofactors for 2nd randomization so must be done 1st.

4. Same end
a) Independent randomizations: Order does not matter



- All combinations of levels of the factors from the two randomized tiers occur.
- There is no confounding of effects from the two randomized tiers.

Decomposition for independent randomizations

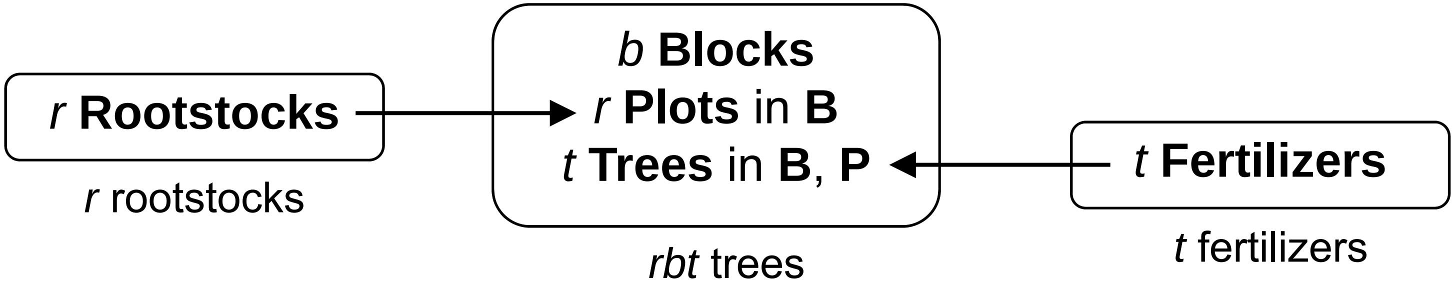
$$\begin{array}{ccccc} \Gamma & \longrightarrow & \Omega & \longleftarrow & \Upsilon \\ \mathcal{T} & & \mathcal{U} & & \mathcal{R} \end{array}$$

- Two structurally balanced designs ($\mathcal{T}$ and $\mathcal{R}$ are in relation to $\mathcal{U}$) are chosen so that for any **U**, except the Mean:
  - either all **UT** are zero,
  - or all **UR** are zero.
- Then $\mathcal{T}$ and $\mathcal{R}$ are independent of each other as only one occurs in any **U**.
- Theorem: given above properties
  - $\mathcal{T}$ is structure balanced in relation to $\mathcal{U} \triangleright \mathcal{R}$;
  - $\mathcal{R}$ is structure balanced in relation to $\mathcal{U} \triangleright \mathcal{T}$;
  - $(\mathcal{U} \triangleright \mathcal{R}) \triangleright \mathcal{T} = (\mathcal{U} \triangleright \mathcal{T}) \triangleright \mathcal{R}$.

Superimposed experiment using split-plots

(Brien & Bailey, 2006, Example 6)

- An RCBD with *b* blocks is set up to investigate the yield differences between *r* rootstocks for orange trees, each plot containing *t* trees.
- After several years of running this initial experiment, decide to incorporate *t* fertilizer treatments by randomizing them to the *t* trees in each plot.



- Two randomizations as they are separated in time and so done separately

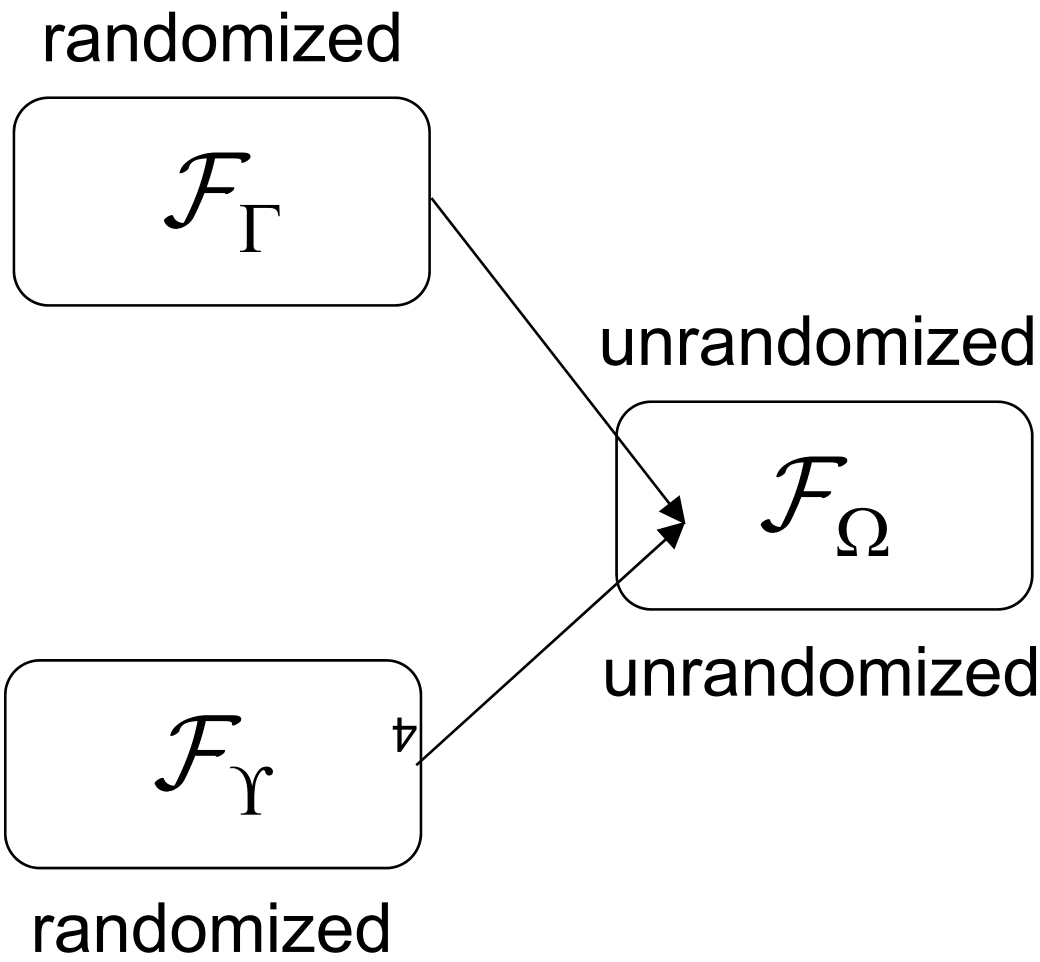
Idempotents for superimposed split-plot

trees tier		rootstocks tier		fertilizers tier	
source	$\mathcal{U}$	source	$\mathcal{U} \rhd \mathcal{R}$	source	$\mathcal{U} \rhd \mathcal{R} \rhd \mathcal{T}$
Mean	$\mathbf{U}_{\mathrm{Mean}}$	Mean	$\mathbf{U}_{\mathrm{Mean}} \rhd \mathbf{R}_{\mathrm{Mean}}$	Mean	$\mathbf{U}_{\mathrm{Mean}} \rhd \mathbf{R}_{\mathrm{Mean}} \rhd \mathbf{T}_{\mathrm{Mean}}$
Blocks	$\mathbf{U}_{\mathrm{B}}$				
Plots[B]	$\mathbf{U}_{\mathrm{P[B]}}$	Rootstocks	$\mathbf{U}_{\mathrm{P[B]}} \rhd \mathbf{R}_{\mathrm{R}} = \mathbf{R}_{\mathrm{R}}$		
		Residual	$\mathbf{U}_{\mathrm{P[B]}} \vdash \mathcal{R}$		
Trees[B∧P]	$\mathbf{U}_{\mathrm{T[B \wedge P]}}$			Fertilizer	$\mathbf{U}_{\mathrm{T[B \wedge P]}} \rhd \mathbf{T}_{\mathrm{F}} = \mathbf{T}_{\mathrm{F}}$
				Residual	$\mathbf{U}_{\mathrm{T[B \wedge P]}} \vdash \mathcal{T}$

$\mathcal{R} = \{\mathbf{R}_{\mathrm{Mean}}, \mathbf{R}_{\mathrm{R}}\}$ $\mathcal{T} = \{\mathbf{T}_{\mathrm{Mean}}, \mathbf{T}_{\mathrm{F}}\}$

- In this case
  - all **T** are orthogonal to $\mathbf{U}_{\mathrm{P[B]}}$;
  - all **R** are orthogonal to $\mathbf{U}_{\mathrm{T[B \wedge P]}}$;

4. Same end
b) Coincident randomizations: Order does not matter



- Levels of some factors from the two randomized tiers are associated by randomization.
- Some effect from one randomized tier is confounded with some effect from the other randomized tier.

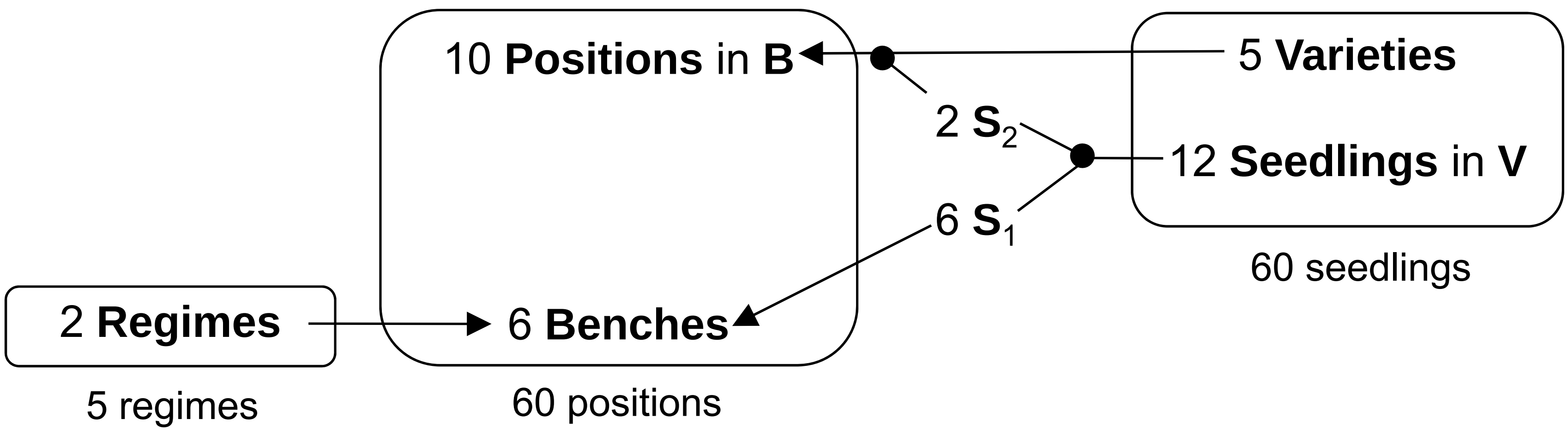
Decomposition for coincident randomizations

$$\begin{array}{ccccc} \Gamma & \longrightarrow & \Omega & \longleftarrow & \Upsilon \\ \mathcal{T} & & \mathcal{U} & & \mathcal{R} \end{array}$$

- Two structurally balanced designs are chosen so that for all **U**, **R** and **T**:
  - either **UR** is zero, (in **U**, ignore **R**)
  - or **UT** is zero (in **U**, ignore **T**)
  - or **U**▷**R** = **U**, (in **U**, do **R** before **T**)
  - or **U**▷**T** = **U**. (in **U**, do **T** before **R**)
- Theorem: Given a pair of structurally balanced randomizatons that are independent or coincident the decompositions $\mathcal{U} \rhd \mathcal{R}$ and $\mathcal{U} \rhd \mathcal{T}$ are compatible in the sense that if **A**$\in \mathcal{U} \rhd \mathcal{R}$ and **B**$\in \mathcal{U} \rhd \mathcal{T}$ then **AB** = **BA**.
- Hence, a decomposition of V_Ω is given by
  - $(\mathcal{U} \rhd \mathcal{R}) \square (\mathcal{U} \rhd \mathcal{T})$ = {**AB**: **A**$\in \mathcal{U} \rhd \mathcal{R}$, **B**$\in \mathcal{U} \rhd \mathcal{T}$}
- This is equivalent to $(\mathcal{U} \rhd \mathcal{R}) \rhd \mathcal{T}$ for independent and many coincident randomizations.

A plant experiment (Brien & Bailey, 2006, Example 5)

- 12 seedlings of each of 5 varieties are put into individual pots;
- these 60 seedlings are randomly assigned to 6 benches in such a way that there are 2 seedlings of each variety on each bench.
- 2 spray regimes are randomly assigned to the benches so that each is applied to the pots on 3 benches.



- Note the two arrows to Benches.
- Variety replication has caused the coincident randomization.

Idempotents for a plant experiment

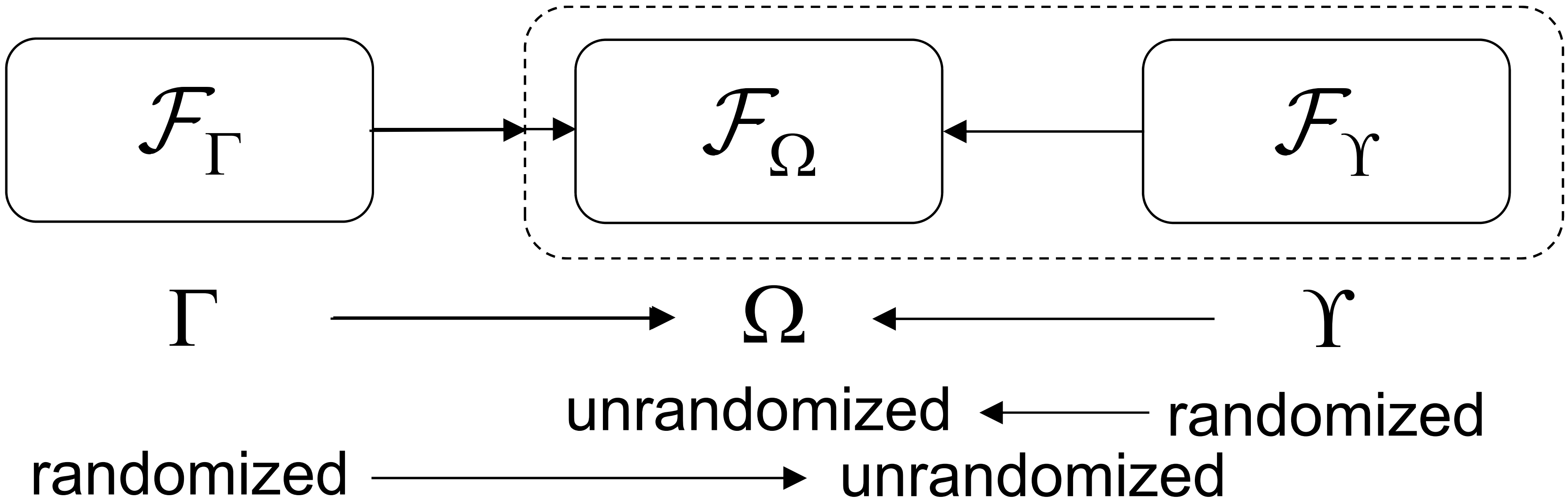
positions tier		seedlings tier		regimes tier	
source	$\mathcal{U}$	source	$\mathcal{U}\rhd\mathcal{R}$	source	$\mathcal{U}\rhd\mathcal{R}\rhd\mathcal{T}$
Mean	$\mathbf{U}_{\mathrm{Mean}}$	Mean	$\mathbf{U}_{\mathrm{Mean}}\rhd\mathbf{R}_{\mathrm{Mean}}$	Mean	$\mathbf{U}_{\mathrm{Mean}}\rhd\mathbf{R}_{\mathrm{Mean}}\rhd\mathbf{T}_{\mathrm{Mean}}$
Benches	$\mathbf{U}_{\mathrm{B}}$	S_1	$\mathbf{U}_{\mathrm{B}}\rhd\mathbf{R}_{\mathrm{S1}} = \mathbf{R}_{\mathrm{S1}}$	Regimes	$(\mathbf{U}_{\mathrm{B}}\rhd\mathbf{R}_{\mathrm{S1}})\rhd\mathbf{T}_{\mathrm{R}} = \mathbf{T}_{\mathrm{R}}$
				Residual	$(\mathbf{U}_{\mathrm{B}}\rhd\mathbf{R}_{\mathrm{S1}})\vdash\mathbf{T}_{\mathrm{R}}$
Positions[B]	$\mathbf{U}_{\mathrm{P[B]}}$	Varieties	$\mathbf{U}_{\mathrm{P[B]}}\rhd\mathbf{R}_{\mathrm{V}} = \mathbf{R}_{\mathrm{V}}$		
		Seedlings[V]$\vdash\mathrm{S}_1$	$\mathbf{U}_{\mathrm{P[B]}}\rhd\mathbf{R}_{\mathrm{VS}\vdash\mathrm{S1}}$		

$\mathcal{R} = \{\mathbf{R}_{\mathrm{Mean}}, \mathbf{R}_{\mathrm{V}}, \mathbf{R}_{\mathrm{S1}}, \mathbf{R}_{\mathrm{VS}\vdash\mathrm{S1}}\}$ $\mathcal{T} = \{\mathbf{T}_{\mathrm{Mean}}, \mathbf{T}_{\mathrm{R}}\}$

- The critical relationship in this case is with $\mathbf{U}_{\mathrm{B}}$
  - From $\mathcal{U}\rhd\mathcal{R}$ we have $\mathbf{U}_{\mathrm{B}}\rhd\mathbf{R}_{\mathrm{S1}} = \mathbf{U}_{\mathrm{B}}$.
  - From $\mathcal{U}\rhd\mathcal{T}$ we have that $\mathbf{U}_{\mathrm{B}}\rhd\mathbf{T}_{\mathrm{R}} = \mathbf{T}_{\mathrm{R}}$.
- The element corresponding to these from $(\mathcal{U}\rhd\mathcal{R})\square(\mathcal{U}\rhd\mathcal{T})$ is
  - $(\mathbf{U}_{\mathrm{B}}\rhd\mathbf{R}_{\mathrm{S1}})(\mathbf{U}_{\mathrm{B}}\rhd\mathbf{T}_{\mathrm{R}}) = \mathbf{U}_{\mathrm{B}}\mathbf{T}_{\mathrm{R}} = \mathbf{T}_{\mathrm{R}}$
- Now $(\mathbf{U}_{\mathrm{B}}\rhd\mathbf{R}_{\mathrm{S1}})(\mathbf{U}_{\mathrm{B}}\rhd\mathbf{T}_{\mathrm{R}}) = (\mathbf{U}_{\mathrm{B}}\rhd\mathbf{R}_{\mathrm{S1}})\rhd\mathbf{T}_{\mathrm{R}}$
- and $(\mathcal{U}\rhd\mathcal{R})\square(\mathcal{U}\rhd\mathcal{T}) = (\mathcal{U}\rhd\mathcal{R})\rhd\mathcal{T}$

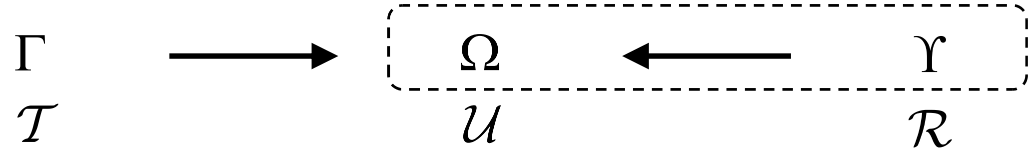
4. Same end

c) Unrandomized-inclusive randomizations: Order does matter



- Two tiers from 1st randomization are unrandomized and form a pseudotier
  - Factors have same status in THIS randomization only.
- Need info from **both** tiers from 1st when doing 2nd randomization.
  - Design and randomization constrained by result of 1st.

Decomposition for unrandomized-inclusive randomizations



- Know $\mathcal{U} \rhd \mathcal{R}$ when randomize Γ.
- Assume:
  - $\mathcal{R}$ is structure balanced in relation to $\mathcal{U}$;
  - $\mathcal{T}$ is structure balanced in relation to $\mathcal{U} \rhd \mathcal{R}$.
- Here must use $(\mathcal{U} \rhd \mathcal{R}) \rhd \mathcal{T}$.
- $\mathcal{U} \rhd \mathcal{R}$ and $\mathcal{U} \rhd \mathcal{T}$ will be compatible for unrandomized-inclusive randomizations if $V_{\Gamma} \cap Mean^{\perp}$ is orthogonal to $V_{\Upsilon} \cap Mean^{\perp}$.

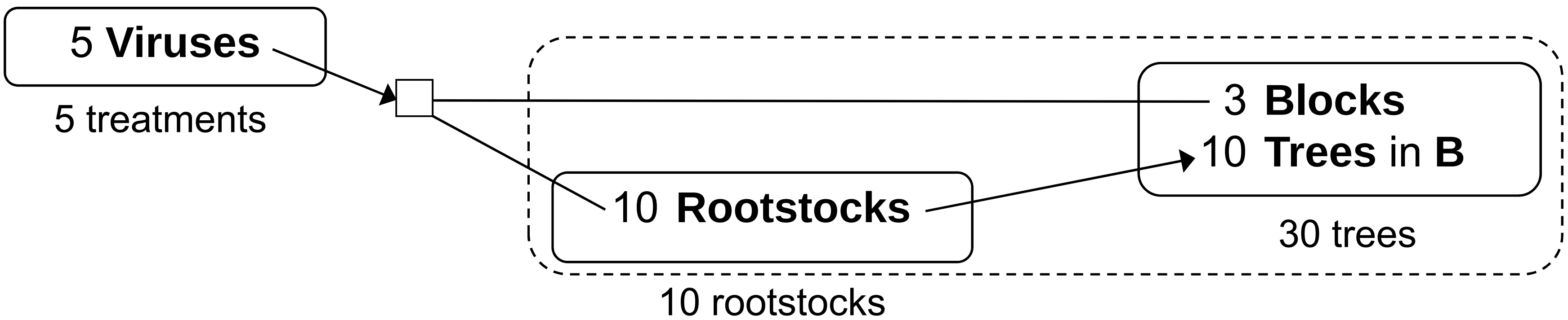
Superimposed Experiment Using a Row-Column Design (Brien & Bailey, 2006, Example 10)

(Freeman, 1959)

- A cherry experiment:
  - originally an RCBD with 3 blocks of 10 rootstocks for 20 years;
  - assign 5 virus treatments to block-rootstock combinations using a balanced row-column design.

		Rootstocks									
		1	2	3	4	5	6	7	8	9	10
	I	A	B	A	C	D	C	B	E	E	D
Blocks	II	D	E	B	D	E	A	C	C	A	B
	III	E	A	C	E	B	D	D	B	C	A

U-inclusive randomizations in example



- Involves 2 randomizations: rootstocks to plots; treatments to trees, taking account of rootstocks.
- Design and randomization of 2nd constrained by result of 1st.
- **Unrandomized** factors for 2nd randomization **inclusive** of factors from both tiers of 1st
  $\Rightarrow$ unrandomized-inclusive randomizations.

Idempotents for cherry experiment

trees tier		rootstocks tier		treatments tier		
source	$\mathcal{U}$	source	$\mathcal{U}\rhd\mathcal{R}$	eff	source	$(\mathcal{U}\rhd\mathcal{R})\rhd\mathcal{T}$
Mean	$\mathbf{U}_{\mathrm{Mean}}$	Mean	$\mathbf{U}_{\mathrm{Mean}}\rhd\mathbf{R}_{\mathrm{Mean}}$		Mean	$\mathbf{T}_{\mathrm{Mean}}$
Blocks	$\mathbf{U}_{\mathrm{B}}$					
Trees[B]	$\mathbf{U}_{\mathrm{T[B]}}$	Rootstocks	$\mathbf{U}_{\mathrm{T[B]}}\rhd\mathbf{R}_{\mathrm{R}} = \mathbf{R}_{\mathrm{R}}$	1/6	Viruses	$(\mathbf{U}_{\mathrm{T[B]}}\rhd\mathbf{R}_{\mathrm{R}})\rhd\mathbf{T}_{\mathrm{V}}$
					Residual	$(\mathbf{U}_{\mathrm{T[B]}}\rhd\mathbf{R}_{\mathrm{R}})\vdash\mathcal{T}$
		Residual	$\mathbf{U}_{\mathrm{T[B]}}\vdash\mathcal{R}$	5/6	Viruses	$(\mathbf{U}_{\mathrm{T[B]}}\vdash\mathcal{R})\rhd\mathbf{T}_{\mathrm{V}}$
					Residual	$(\mathbf{U}_{\mathrm{T[B]}}\vdash\mathcal{R})\vdash\mathbf{T}_{\mathrm{V}}$

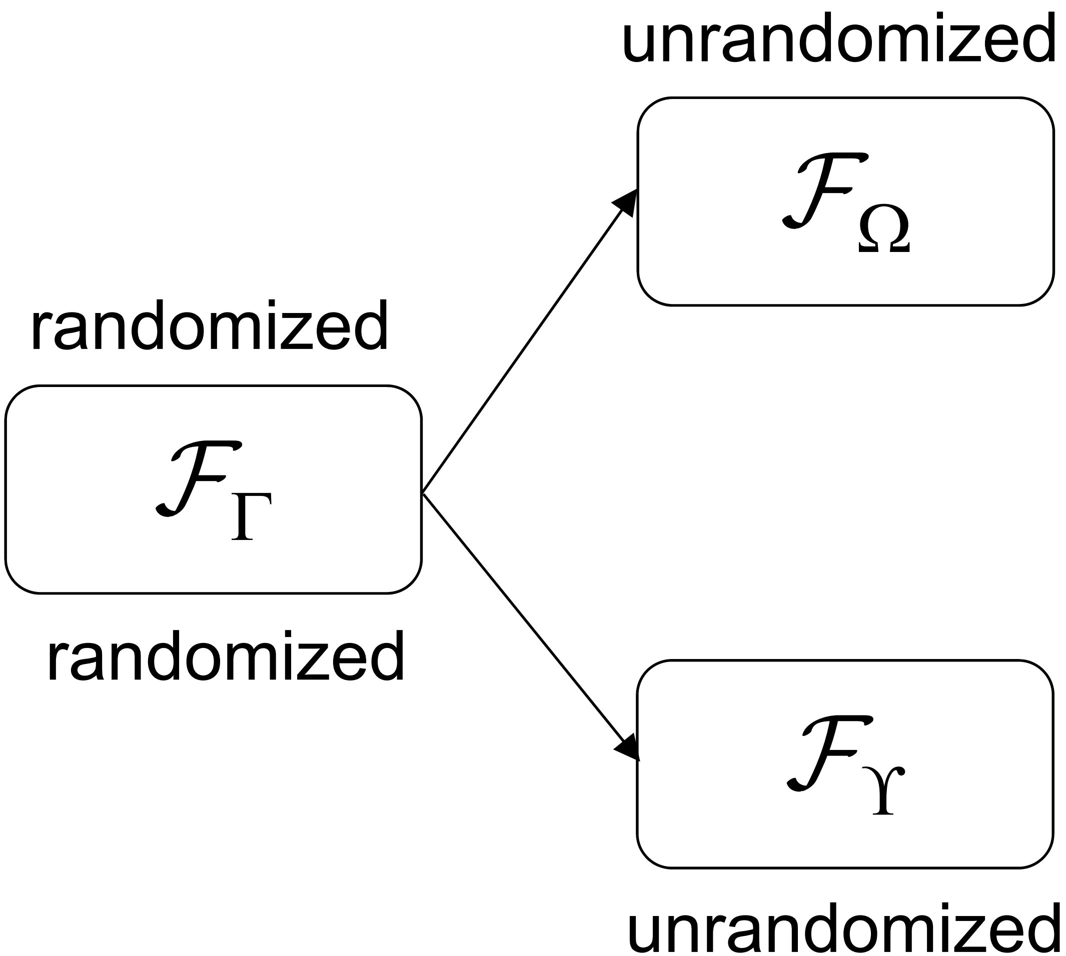
$\mathcal{R} = \{\mathbf{R}_{\mathrm{Mean}}, \mathbf{R}_{\mathrm{R}}\}$ $\qquad$ $\mathcal{T} = \{\mathbf{T}_{\mathrm{Mean}}, \mathbf{T}_{\mathrm{V}}\}$

- Theorem: If
  - $\mathcal{R}$ is structure balanced in relation to $\mathcal{U}$, and
  - $\mathcal{T}$ is structure balanced in relation to $\mathcal{U}\rhd\mathcal{R}$
- Then $\mathcal{T}$ is structure balanced in relation to $\mathcal{U}$, with efficiency factors
  - $\lambda_{\mathbf{UT}} = \lambda_{\mathbf{U}\vdash\mathcal{R},\mathbf{T}} + \sum_{\mathbf{R}} \lambda_{\mathbf{U}\rhd\mathbf{R},\mathbf{T}}$.
- For the example we have $\lambda_{\mathrm{T[B],V}} = 5/6 + 1/6 = 1$.

Incoherent unrandomized-inclusive randomizations (Brien & Bailey, 2006, section 5.2.1)

- For u-inclusive randomizations, factors nested in first randomization can be crossed in second.
  - In Example 10:
    - for 1st randomization, Trees nested in Blocks;
    - for 2nd randomization, Rootstocks replaces Trees and Rootstocks and Blocks crossed.
- But, factors crossed in 1st randomization must be crossed in 2nd randomization.
  - If not, say the randomizations are **incoherent**.
  - If in example 10 Viruses randomized to Blocks within Rootstocks using a BIBD, then incoherent:
    - Blocks thought important in 1st randomization but overridden in 2nd randomization.
- Example 11, a widely advocated split-plot design, is a more complicated case of incoherent multiple randomizations.

5. Same start
Double randomizations:
Order does not matter

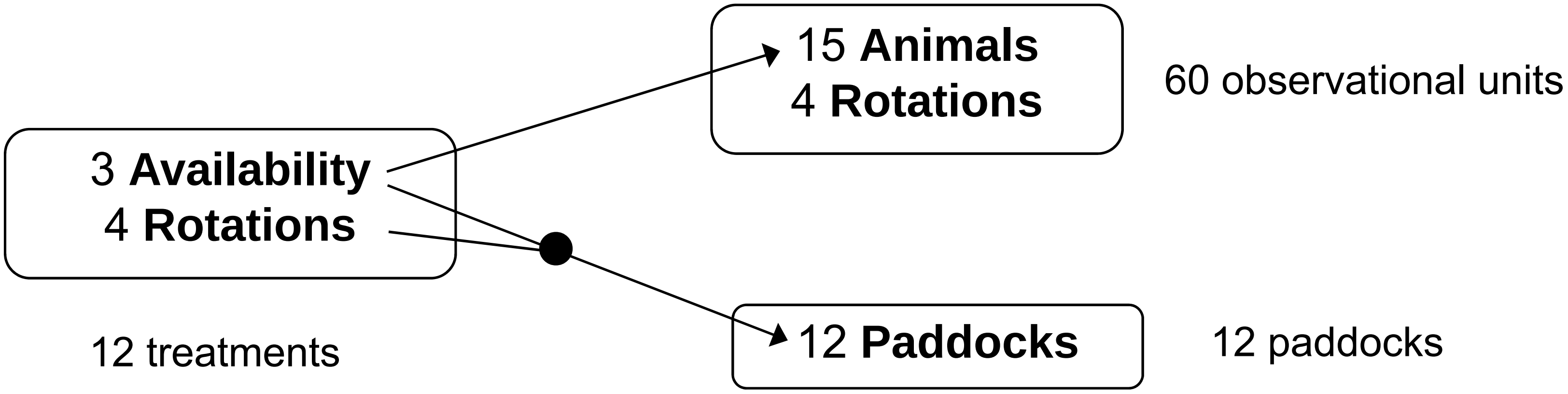


- One unrandomized set has the same size as the doubly randomized set; the other contains the observational units.
- Degenerate case of randomized-inclusive randomization.

An improperly replicated grazing experiment

(Brien & Bailey, 2006, Example 8)

- Combinations of 3 levels of availability and 4 rotations are applied completely at random to 12 paddocks.
- Also, the levels of availability are assigned completely at random to 15 animals so that each level of availability is assigned to 5 animals.
- The 5 animals are then grazed together in sequence on the 4 paddocks assigned to that level of availability;
  - the sequence of 4 paddocks is determined by the rotations assigned to them.



6. Designing multitiered experiments

6(i) Two randomizations

(Brien and Bailey, Section 8.1)

Chain randomizations

a) *Composed* is used where all information from each randomized source is to be confounded with the same source(s) and neither randomization needs information from the outcome of the other.
   - it occurs in two-phase experiments and grazing trials.

b) *Randomized inclusive* is used where information from a source is to be subdivided according to results of a first randomization and different parts are to be confounded with different sources.
   - it occurs in two-phase and multistage reprocessing experiments.

6(i) Two randomizations (cont'd)

Two to one randomizations

c) *Coincident* is used where two sources from different randomizations are to be confounded with the same source(s), and so with each other.

- e.g. when treatments are introduced in a lab phase and replicates from a field phase are to be confounded with the same laboratory phase source;
- it occurs in a range of experiments including two-phase, single-stage and multistage reprocessing experiments.

d) *Independent* is used to randomize sources from different tiers to different sources from the same tier.

- it occurs in a range of experiments including two-phase, superimposed and multistage reprocessing experiments

e) *Unrandomized inclusive* is used where unrandomized factors for the second randomization must take into account the assignment of all factors from the first randomization.

- it occurs in superimposed experiments and two-phase experiments.

One to two randomizations

f) *Double* is a degenerate type

- rotational grazing trials are the only known examples.

6(ii) Three or more randomizations

(Brien and Bailey, Section 6)

- Many possible combinations of the 6 types of multiple randomization.
- Example 12–14 (Brien and Bailey, 2006) are some possibilities.

6(iii) Multitiered experiments in practice

- Not new designs but a particular class of (complex) experiments (see http://chris.brien.name/multitier).
  - Two-phase and multiphase experiments (Brien and Bailey, 2006, Examples 1, 4, 9, 12, 13, 14, 15, Figure 7; Wilkinson et al, 2008);
  - Superimposed experiments (Brien and Bailey, 2006, Examples 6, 10);
  - Plant experiments (McIntyre, 1955; Preece, 1991; Brien and Bailey, 2006, Example 5);
  - Grazing experiments (Brien and Demétro, 1998; Brien and Bailey, 2006, Examples 3, 8);
  - Human interaction experiments (Lewis and Russell, 1998 ; Brien and Bailey, 2006, Figure 28);
  - Multistage reprocessing experiments (Miller, 1997; Mee and Bates, 1998).
- Of these, two-phase experiments are perhaps the most commonly occurring, although often unrecognized.

6(iv) Second phase a lab phase

- Important and widely occurring, but unrecognized, subclass with second phase a laboratory phase:
  - Field + lab phases;
  - Glasshouse + lab phases;
  - Field + sensory phases;
  - Food preparation and processing phases;
  - Biological expt + microarray expt;
  - Clinical trial + lab phase.
- Generally randomization of the lab phase not considered
  - Ignores possibility of systematic trends in the laboratory phase
  - Ad hoc evidence that systematic trend may be widespread
- If do, a randomization to different units, in each phase, is required

  $\Rightarrow$ Multiple randomizations

Lab phase design

- Similar to standard design: randomize a bunch of factors to another bunch — just more factors
- Lab units often sequence of analyses in time (and/or space) and how should one group these and assign 1st phase units?
- Important to have some idea of likely laboratory variation:
  - Will there be recalibration or the like?
  - Are consistent differences between and/or across Occasions likely?
  - How does the magnitude of the field and laboratory variation compare?
  - Are trends probable: common vs different; linear vs cubic?
- Will laboratory duplicates be necessary and how will they be arranged?
- Treatments to be added in lab phase: leads to more randomizations (e.g. corn experiment)
- Cannot improve on field design but can make worse.

Randomizing the lab phase

- Composed or randomized-inclusive related to whether 1st – phase, unrandomized factors match lab, unrandomized factors.
- Fundamental difference between 1st and 2nd randomizations:
  - 1st has single smaller set of randomized (treatment) factors with all levels combinations observable (cf. single randomization);
  - 2nd has two sets of factors to be randomized, but not all levels combinations of the factors observable
    $\Rightarrow$ tendency to ignore 1st phase unit factors.
- Approach does not omit 1st-phase factors
  - uses pseudofactors with randomzed-inclusive to avoid this and make explicit what has occurred.
- Categories of designs
  - Lab phase factors *purely hierarchical* or *involve crossed rows and columns*;
  - Two-phase randomizations are *composed* or *randomized-inclusive*;
  - *Treatments added in laboratory phase* or *not*;
  - *Lab duplicates included* or *not.*

7. Pseudofactors

- Pseudofactors (Yates, 1936; Monod and Bailey, 1992) group the levels of a factor with the feature that there is no scientific rationale for the grouping.
  - done as an aid in the design and analysis of the experiment.
- Occur more often in multitiered experiments
- Mechanisms:
  a) Purposefully-chosen groups so resulting subspace has certain desirable properties.
     - The L-pseudofactors for Lines in the wheat experiment.

  b) Grouped using factors randomized to the factor from which the pseudofactor is derived.
     - The P-pseudofactors for Plots in the wheat experiment, where the Lines, that are randomized to the Plots, determine the pseudofactors.

  c) Randomly-chosen groups, deliberately avoiding systematic bias.
     - The S-pseudofactors for Seedlings in the plant experiment.

8. Summary

- Multitiered experiments involve multiple randomizations and so more than two tiers.
- They are widely applicable.
- Six types of multiple randomizations that can be classified according to:
  - pattern in arrows: chain, same end, same start;
  - whether or not order of two randomizations matter.
- The types differ in:
  - the forms of idempotents that can occur;
  - in the properties of their efficiency factors e.g. efficiency factors for chain randomizations multiply.
- Pseudofactors occur more often in multitiered experiments.
  - Three mechanisms for generating them.