VI.E Determining the analysis of variance table – further examples

For each of the following studies determine the analysis of variance table (Source, df, SSq and E[MSq]) using the following seven steps:

a) Description of pertinent features of the study
b) The experimental structure
c) Sources derived from the structure formulae
d) Degrees of freedom and sums of squares
e) The analysis of variance table
f) Maximal expectation and variation models
g) The expected mean squares.

Example VI.8 Mathematics teaching methods

An educational psychologist wants to determine the effect of three different methods of teaching mathematics to year 10 students. Five metropolitan schools with three mathematics classes in year 10 are selected and the methods of teaching randomized to the classes in each school. After being taught by one of the methods for a semester, the students sit a test and their average score is recorded.

a) Description of pertinent features of the study

1. the observational unit – a class
2. response variable – Test score
3. unrandomized factors – Schools, Classes
4. randomized factors – Methods
5. type of study – an RCBD

b) The experimental structure

Structure	Formula
unrandomized	*5* Schools/*3* Classes
randomized	*3* Methods

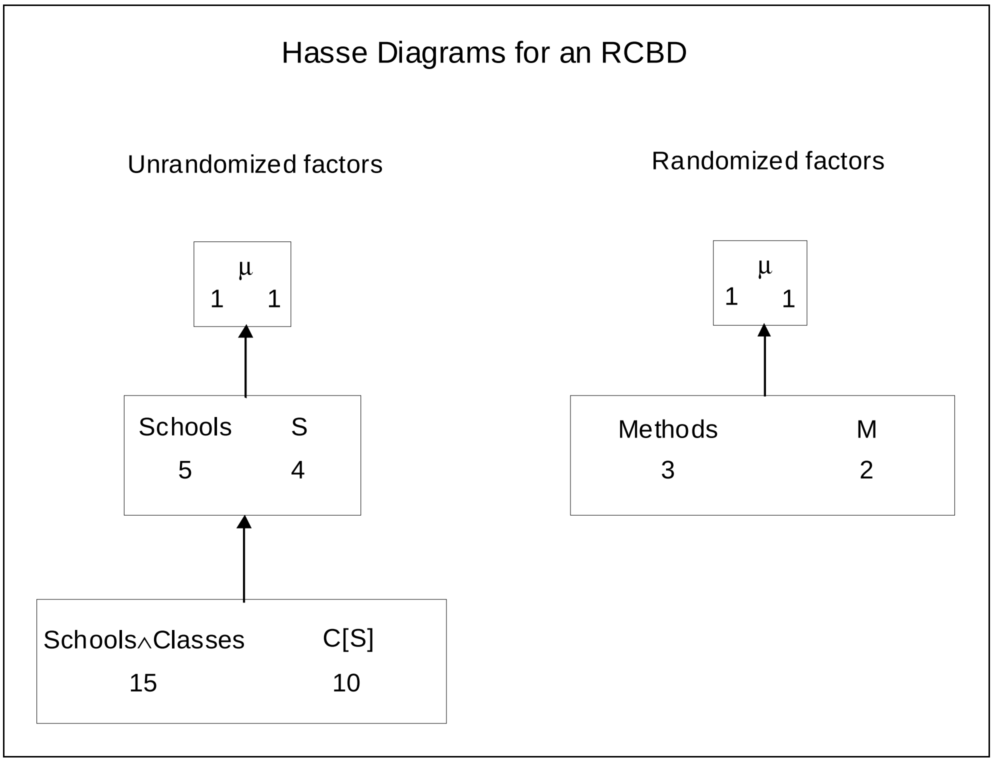
c) Sources derived from the structure formulae

Schools/Classes = Schools + Classes[Schools]

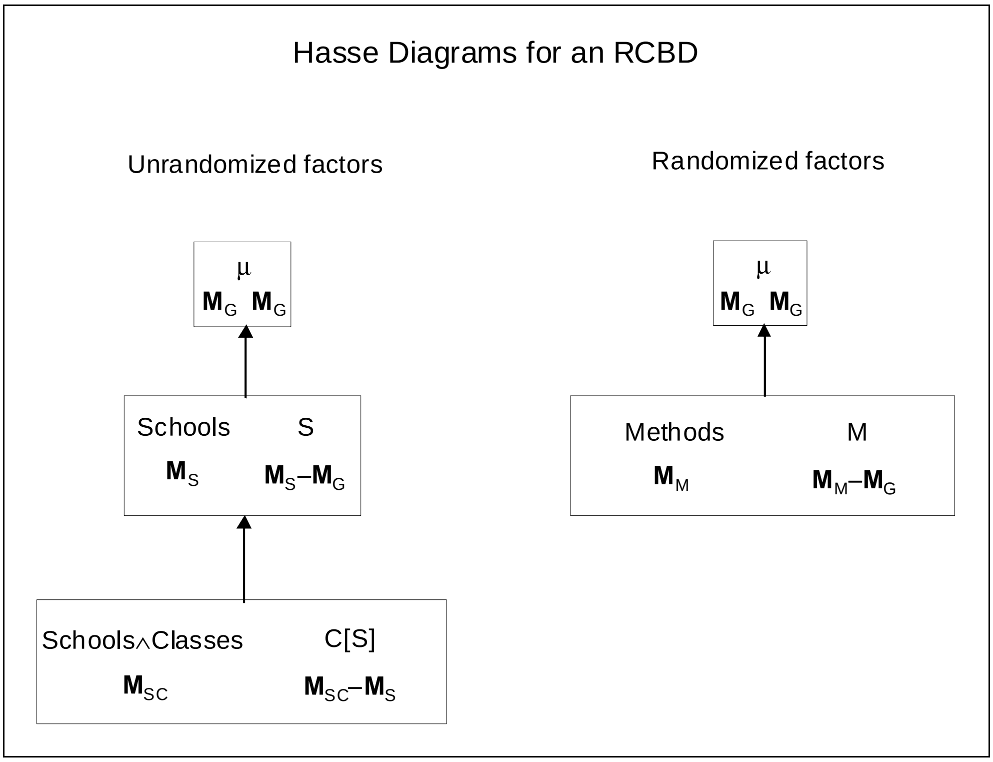
Methods = Methods

d) Degrees of freedom and sums of squares

The Hasse diagrams, with degrees of freedom, for this study are:



The Hasse diagrams, with **M** and **Q** matrices, for this study are:



e) *The analysis of variance table*

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

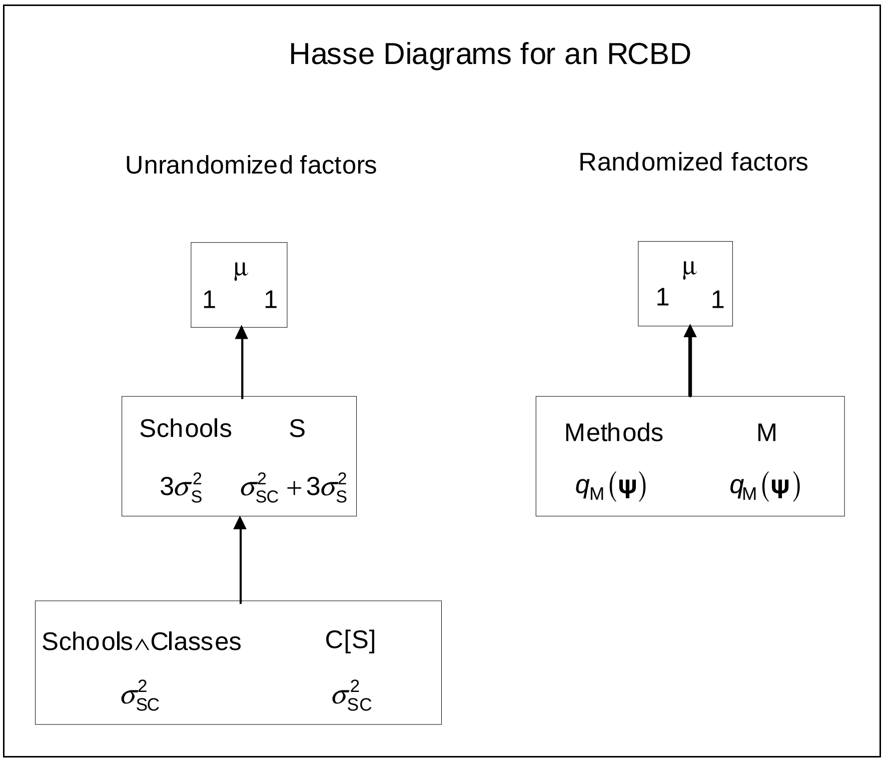
f) Maximal expectation and variation models

Seems that randomized factors should be fixed and unrandomized factors should be random. Hence, the maximal variation and expectation models are:

$$\text{Var}[\mathbf{Y}] = \text{Schools} + \text{Schools}\wedge\text{Classes}$$
$$\psi = \text{E}[\mathbf{Y}] = \text{Methods}$$

g) The expected mean squares.

The Hasse diagrams, with contributions to expected mean squares, for this study are:



The analysis of variance table is:

Source	df	SSq	E[MSq]
Schools	4	$\mathbf{Y}'\mathbf{Q}_{\text{S}}\mathbf{Y}$	$\sigma^2_{\text{SC}} + 3\sigma^2_{\text{S}}$
Classes[Schools]	10	$\mathbf{Y}'\mathbf{Q}_{\text{SC}}\mathbf{Y}$	
Methods	2	$\mathbf{Y}'\mathbf{Q}_{\text{M}}\mathbf{Y}$	$\sigma^2_{\text{SC}} + q_{\text{M}}(\boldsymbol{\psi})$
Residual	8	$\mathbf{Y}'\mathbf{Q}_{\text{SC}_{\text{Res}}}\mathbf{Y}$	σ^2_{SC}

■

Example VI.9 Smoke emission from combustion engines

A manufacturer of combustion engines is concerned about the percentage of smoke emitted by engines of a particular design in order to meet air pollution standards. An experiment was conducted involving engines with three timing levels, three throat diameters, two volume ratios in the combustion chamber, and two injection systems. Thirty-six engines were designed to represent all possible combinations of these four factors and the engines were then operated in a completely random order so that each engine was tested twice. The percent smoke emitted was recorded at each test.

a) Description of pertinent features of the study

1. Observational unit – a test
2. Response variable – Percent Smoke Emitted
3. Unrandomized factors – Tests
4. Randomized factors – Timing, Diameter, Ratios, Systems
5. Type of study – Complete Four-factor CRD

b) The experimental structure

Structure	Formula
unrandomized	*72* Tests
randomized	*3* Timing**3* Diameter**2* Ratios**2* Systems

c) Sources derived from the structure formulae

Tests = Tests

Timing*Diameter*Ratios*Systems
= Timing + Diameter + Ratios + Systems
+ Timing#Diameter + Timing#Ratios + Timing#Systems
+ Diameter#Ratios + Diameter#Systems + RatiosSystems
+ Timing#Diameter#Ratios + Timing#Diameter#Systems
+ Timing#Ratios#Systems + Diameter#Ratios#Systems
+ Timing#Diameter#Ratios#Systems

d) Degrees of freedom and sums of squares

Note that factors in the randomized structure are completely crossed so that the degrees of freedom of a source from that structure can be obtained by computing the number of levels minus one for each factor in the source and forming their product.

e) The analysis of variance table

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

f) Maximal expectation and variation models

Seems that randomized factors should be fixed and unrandomized factor should be random. Hence, the maximal variation and expectation models are:

$$\text{Var}[Y] = \text{Tests}$$
$$\psi = \text{E}[Y] = \text{Timing}\wedge\text{Diameter}\wedge\text{Ratios}\wedge\text{Systems}$$

The expectation model consists of just the generalized factor corresponding to the source Timing#Diameter#Ratios#Systems; all other expectation generalized factors are marginal to this generalized factor.

g) The expected mean squares.

In this case, using the Hasse diagrams to compute the E[MSq]s would be overkill. There is one variation component for Tests: $(72/72)\sigma_t^2 = \sigma_t^2$. Since Tests is partitioned into the other sources in the table, this component occurs against all lines except Tests. All the randomized generalized factors are potential contributors to the expectation model and so their contributions to the E[MSq]s for the corresponding sources are of the form $q_F(\boldsymbol{\psi})$.

The analysis of variance table is:

Source	df	SSq	E[MSq]
Tests	71	$\mathbf{Y}'\mathbf{Q}_t\mathbf{Y}$	
Timing	2	$\mathbf{Y}'\mathbf{Q}_T\mathbf{Y}$	$\sigma_t^2 + q_T(\boldsymbol{\psi})$
Diameter	2	$\mathbf{Y}'\mathbf{Q}_D\mathbf{Y}$	$\sigma_t^2 + q_D(\boldsymbol{\psi})$
Ratios	1	$\mathbf{Y}'\mathbf{Q}_R\mathbf{Y}$	$\sigma_t^2 + q_R(\boldsymbol{\psi})$
Systems	1	$\mathbf{Y}'\mathbf{Q}_S\mathbf{Y}$	$\sigma_t^2 + q_S(\boldsymbol{\psi})$
Timing#Diameter	4	$\mathbf{Y}'\mathbf{Q}_{TD}\mathbf{Y}$	$\sigma_t^2 + q_{TD}(\boldsymbol{\psi})$
Timing#Ratios	2	$\mathbf{Y}'\mathbf{Q}_{TR}\mathbf{Y}$	$\sigma_t^2 + q_{TR}(\boldsymbol{\psi})$
Timing#Systems	2	$\mathbf{Y}'\mathbf{Q}_{TS}\mathbf{Y}$	$\sigma_t^2 + q_{TS}(\boldsymbol{\psi})$
Diameter#Ratios	2	$\mathbf{Y}'\mathbf{Q}_{DR}\mathbf{Y}$	$\sigma_t^2 + q_{DR}(\boldsymbol{\psi})$
Diameter#Systems	2	$\mathbf{Y}'\mathbf{Q}_{DS}\mathbf{Y}$	$\sigma_t^2 + q_{DS}(\boldsymbol{\psi})$
Ratios#Systems	1	$\mathbf{Y}'\mathbf{Q}_{RS}\mathbf{Y}$	$\sigma_t^2 + q_{RS}(\boldsymbol{\psi})$
Timing#Diameter#Ratios	4	$\mathbf{Y}'\mathbf{Q}_{TDR}\mathbf{Y}$	$\sigma_t^2 + q_{TDR}(\boldsymbol{\psi})$
Timing#Diameter#Systems	4	$\mathbf{Y}'\mathbf{Q}_{TDS}\mathbf{Y}$	$\sigma_t^2 + q_{TDS}(\boldsymbol{\psi})$
Timing#Ratios#Systems	2	$\mathbf{Y}'\mathbf{Q}_{TRS}\mathbf{Y}$	$\sigma_t^2 + q_{TRS}(\boldsymbol{\psi})$
Diameter#Ratios#Systems	2	$\mathbf{Y}'\mathbf{Q}_{DRS}\mathbf{Y}$	$\sigma_t^2 + q_{DRS}(\boldsymbol{\psi})$
Timing#Diameter#Ratios#System	4	$\mathbf{Y}'\mathbf{Q}_{TDRS}\mathbf{Y}$	$\sigma_t^2 + q_{TDRS}(\boldsymbol{\psi})$
Residual	36	$\mathbf{Y}'\mathbf{Q}_{t_{Res}}\mathbf{Y}$	σ_t^2

■

Example VI.10 Lead concentration in hair

An investigation was performed to discover trace metal concentrations in humans in the five major cities in Australia. The concentration of lead in the hair of fourth grade school boys was determined. In each city, 10 primary schools were randomly selected and from each school ten students selected. Hair samples were taken from the selected boys and the concentration of lead in the hair determined.

a) Description of pertinent features of the study

1. the observational unit – hair of a boy
2. response variable – Lead concentration
3. unrandomized factors – Cities, Schools, Boys
4. randomized factors – not applicable
5. type of study – a multistage survey

b) The experimental structure

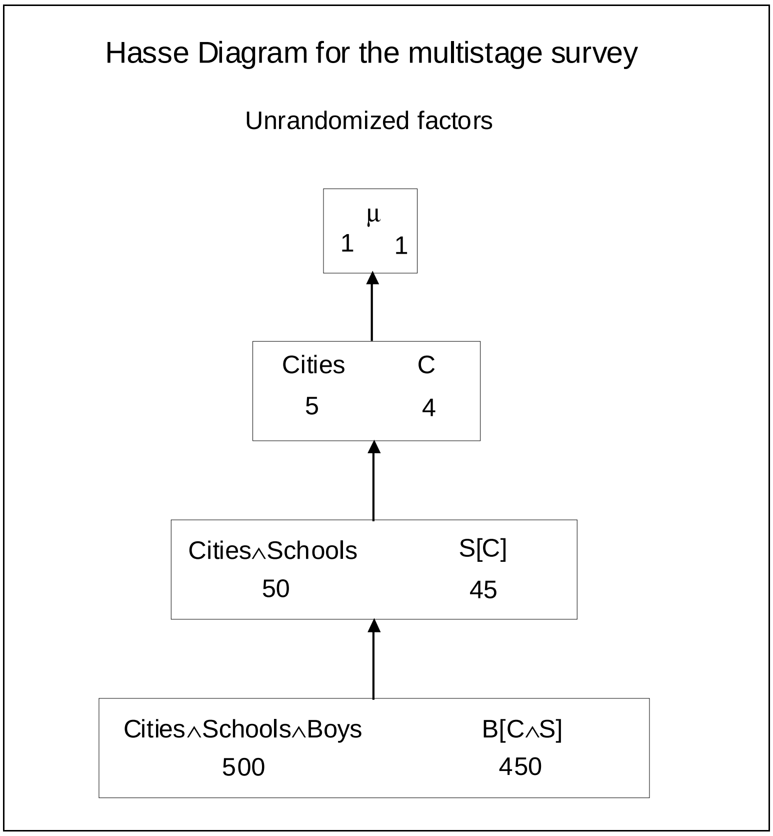
Structure	Formula
unrandomized	*5* Cities/*10* Schools/*10* Boys
randomized	not applicable.

c) Sources derived from the structure formulae

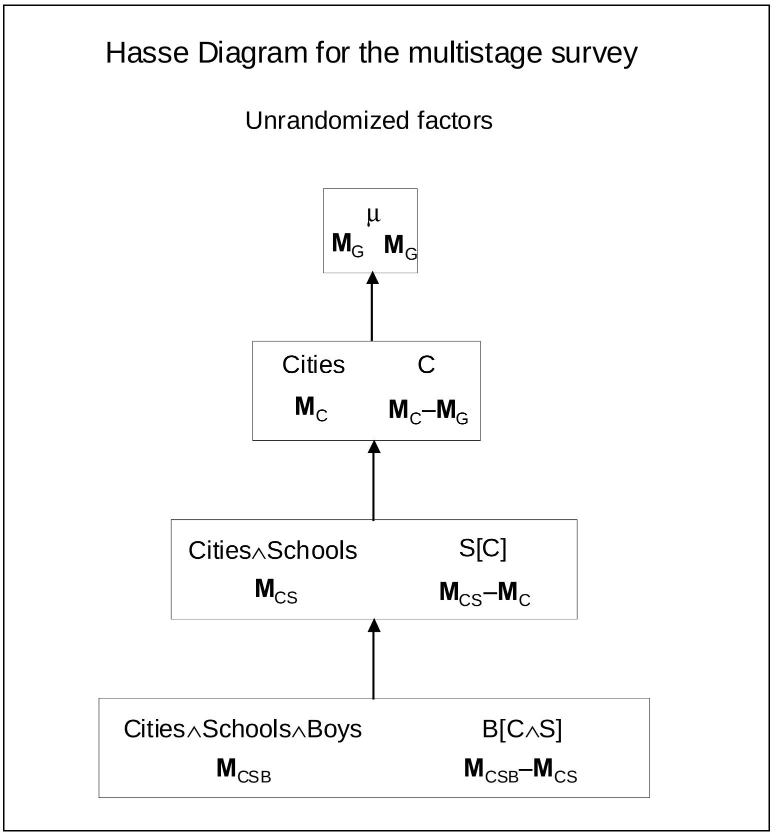
Cities/Schools/Boys = (Cities + Schools[Cities])/Boys
= Cities + Schools[Cities]
+ Boys[*gf*(Cities + Schools[Cities])]
= Cities + Schools[Cities] + Boys[Cities$\wedge$Schools]

d) Degrees of freedom and sums of squares

The Hasse diagram, with degrees of freedom, for the unrandomized generalized factors in this study is:



The Hasse diagram, with **M** and **Q** matrices, for the unrandomized generalized factors in this study is:



e) The analysis of variance table

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

f) Maximal expectation and variation models

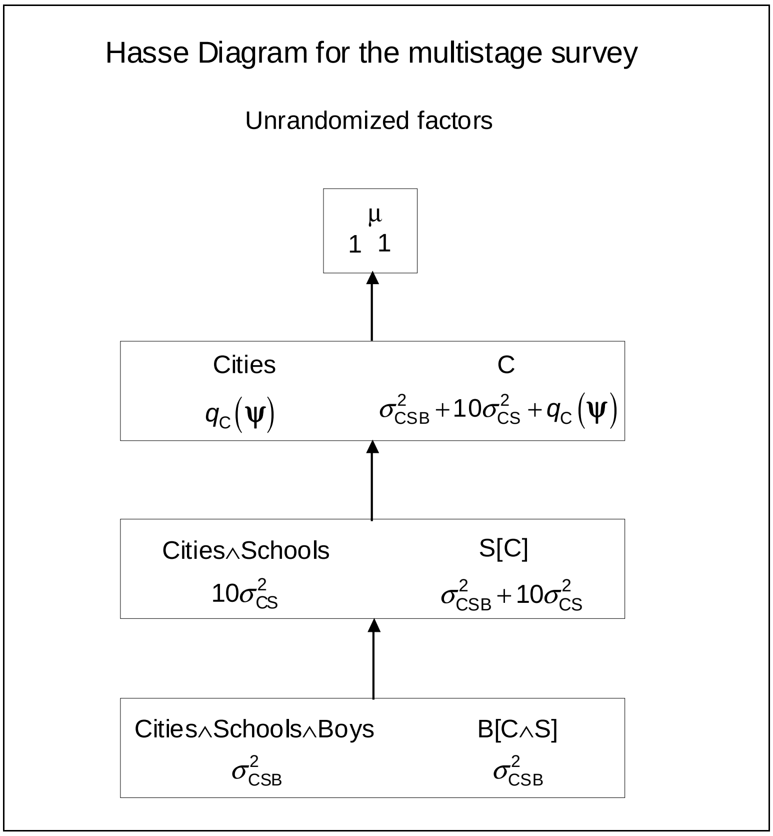
Seems that Cities should be fixed and Schools and Boys should be random. Hence, the maximal variation and expectation models are:

$$\text{Var}[Y] = \text{Cities}\wedge\text{Schools} + \text{Cities}\wedge\text{Schools}\wedge\text{Boys and}$$

$$\psi = \text{E}[Y] = \text{Cities}$$

g) The expected mean squares.

The Hasse diagrams, with contributions to expected mean squares, for this study are:



The analysis of variance table is

Source	df	SSq	E[MSq]
Cities	4	$\mathbf{Y}'\mathbf{Q}_{\mathrm{C}}\mathbf{Y}$	$\sigma^2_{\mathrm{CSB}} + 10\sigma^2_{\mathrm{CS}} + q_{\mathrm{C}}(\boldsymbol{\psi})$
Schools[Cities]	45	$\mathbf{Y}'\mathbf{Q}_{\mathrm{CS}}\mathbf{Y}$	$\sigma^2_{\mathrm{CSB}} + 10\sigma^2_{\mathrm{CS}}$
Boys[Cities$\wedge$Schools]	450	$\mathbf{Y}'\mathbf{Q}_{\mathrm{CSB}}\mathbf{Y}$	σ^2_{CSB}

■

Example VI.11 Penicillin pain

An experiment was conducted to investigate the degree of pain experienced by patients when injected with penicillin of different potencies. In the experiment there were 10 groups of three patients, all three patients being simultaneously injected with the same dose. Altogether there were six different potencies, all of which were administered over six consecutive times to the groups of patients; the order in which they were administered is randomized for each group. The total of the degree of pain experienced by the three patients was obtained, the pain for each patient having been measured on a five-point scale 0–4.

a) Description of pertinent features of the study

1. the observational unit – group of patients at a time
2. response variable – total pain score
3. unrandomized factors – Group, Time
4. randomized factors – Potency
5. type of study – Randomized Complete Block Design

b) The experimental structure

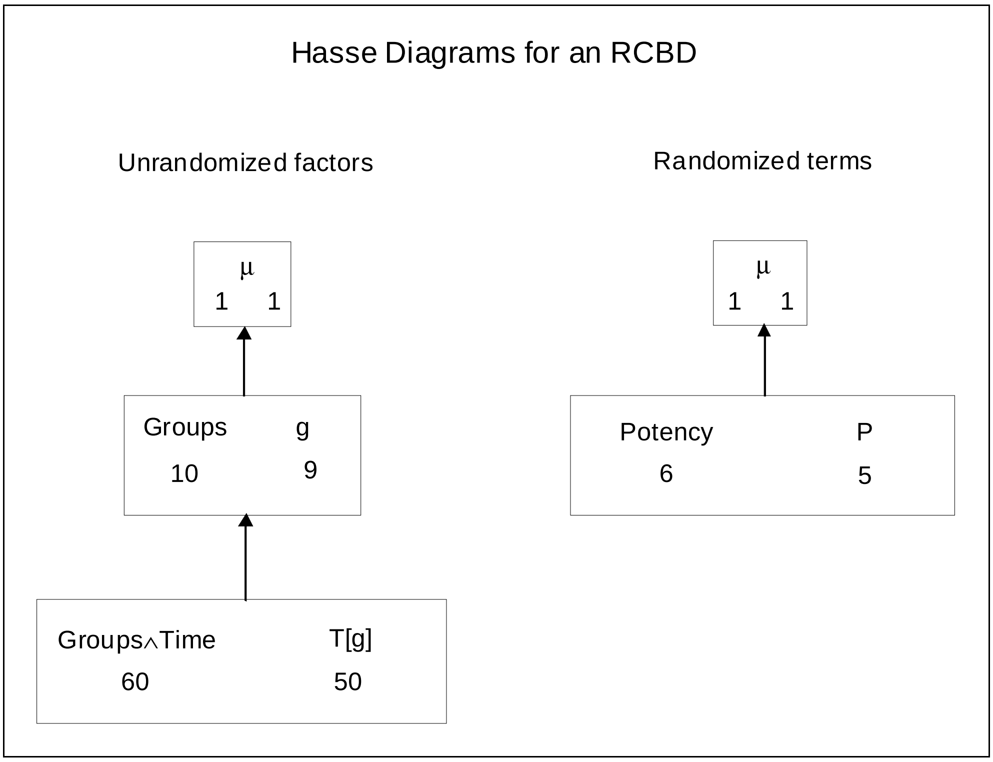
Structure	Formula
unrandomized	*10* Groups/*6* Time
randomized	*6* Potency

c) Sources derived from the structure formulae

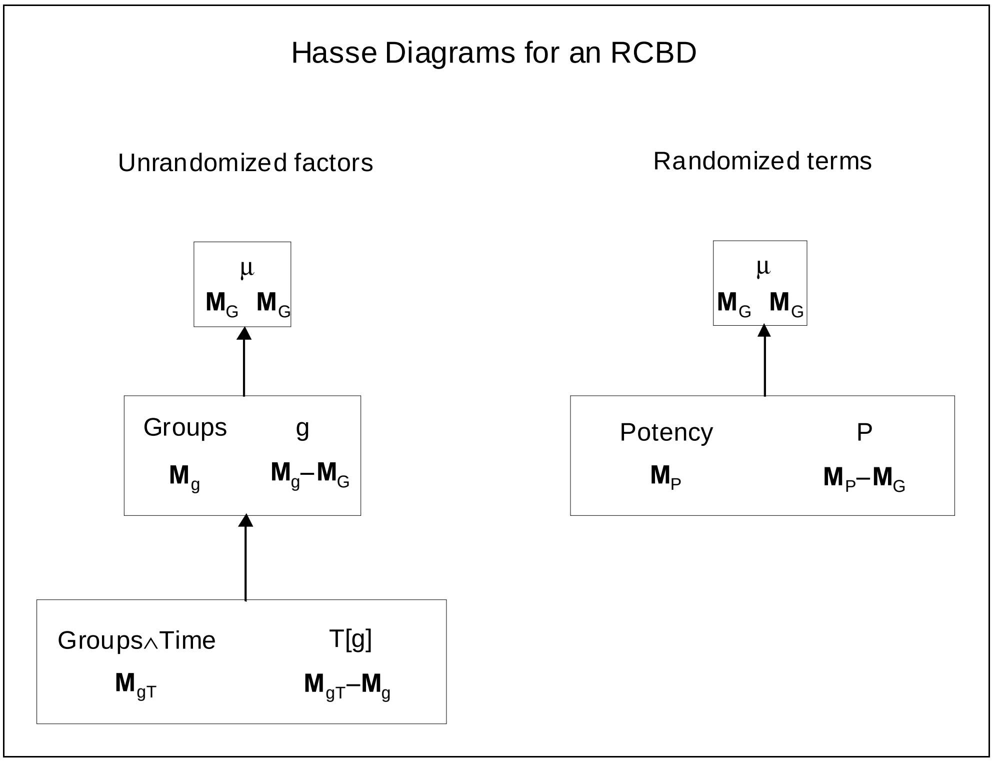
Groups/Time = Groups + Time[Groups]
Potency = Potency

d) Degrees of freedom and sums of squares

The Hasse diagrams, with degrees of freedom, for this study are (use g for Groups to distinguish from G for Grand Mean):



The Hasse diagrams, with **M** and **Q** matrices, for this study are:



e) The analysis of variance table

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

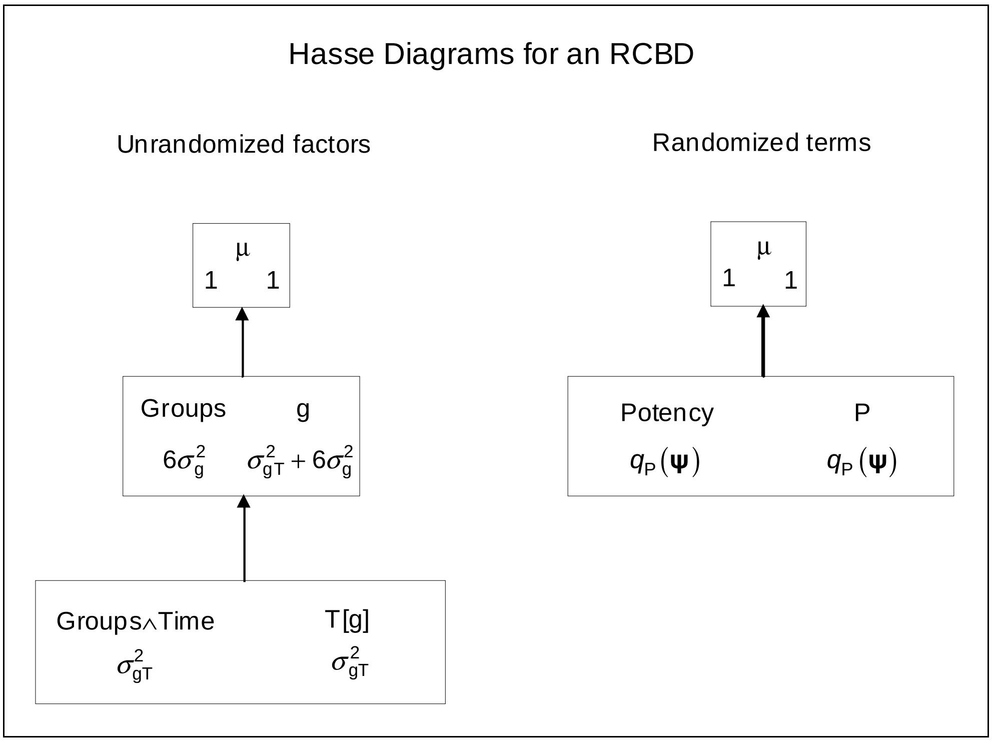
f) Maximal expectation and variation models

Seems that randomized factors should be fixed and unrandomized factors should be random. Hence, the maximal variation and expectation models are:

$$Var[\mathrm{Y}] = \text{Groups} + \text{Groups}\wedge\text{Time}$$
$$\psi = E[\mathrm{Y}] = \text{Potency}$$

g) The expected mean squares.

The Hasse diagrams, with contributions to expected mean squares, for this study are:



The analysis of variance table is:

Source	df	SSq	E[MSq]
Groups	9	$\mathbf{Y}'\mathbf{Q}_{\mathrm{g}}\mathbf{Y}$	$\sigma^2_{\mathrm{gT}} + 6\sigma^2_{\mathrm{g}}$
Time[Groups]	50	$\mathbf{Y}'\mathbf{Q}_{\mathrm{gT}}\mathbf{Y}$	
Potencies	5	$\mathbf{Y}'\mathbf{Q}_{\mathrm{P}}\mathbf{Y}$	$\sigma^2_{\mathrm{gT}} + q_{\mathrm{P}}(\boldsymbol{\psi})$
Residual	45	$\mathbf{Y}'\mathbf{Q}_{\mathrm{gT_{Res}}}\mathbf{Y}$	σ^2_{gT}
Total	59		

■

Example VI.12 Plant rehabilitation study

In a plant rehabilitation study, the increase in height of plants of a certain species during a 12-month period was to be determined at three sites differing in soil salinity. Each site was divided into five parcels of land containing 4 plots and four different management regimes applied to the plots, the regimes being randomized to the plots within a parcel. In each plot, six plants of the species were selected and marked and the total increase in height of all six plants measured.

Note that the researcher is interested in seeing if any regime differences vary from site to site.

a) Description of pertinent features of the study

1. the observational unit – a plot of plants
2. response variable – Plant height increase
3. unrandomized factors – Sites, Parcels, Plots
4. randomized factors – Regimes
5. type of study – a type of RCBD

b) The experimental structure

Structure	Formula
unrandomized	*3* Sites/*5* Parcels/*4* Plots
randomized	*4* Regimes*Sites

Note the unrandomized factor Sites has also been included in the same formula as the randomized factors so that the interaction between Sites and Regimes is investigated, as the researcher has asked.

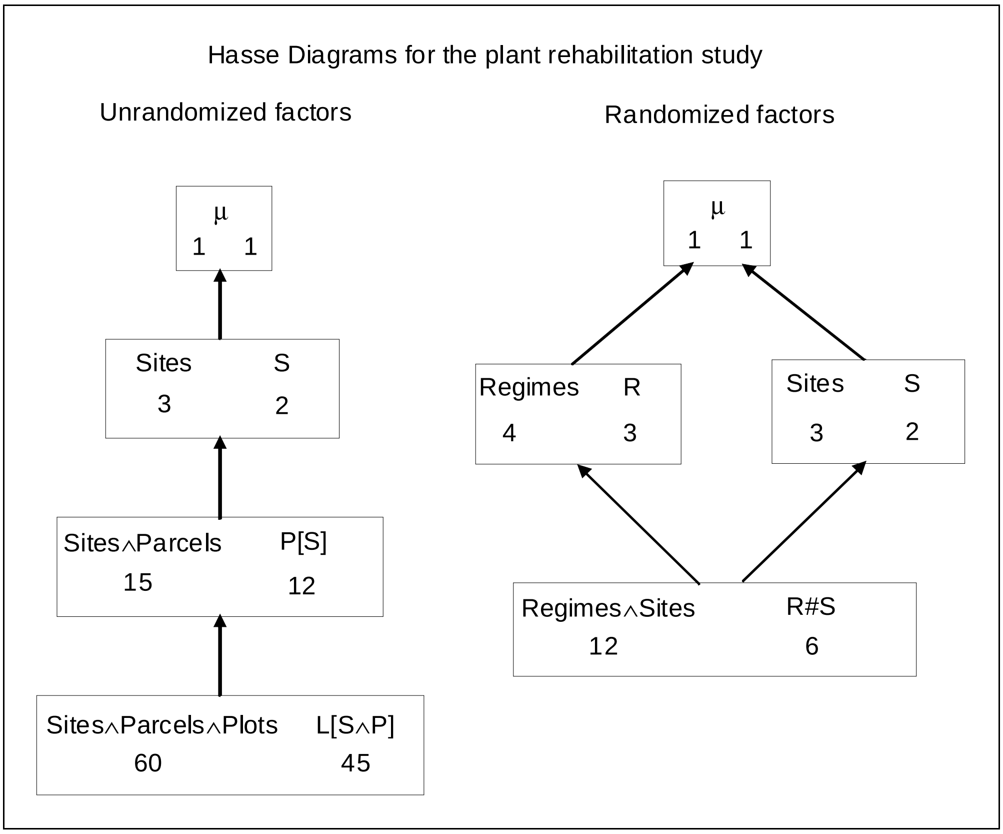
c) Sources derived from the structure formulae

$$\text{Sites/Parcels/Plots} = \text{Sites} + \text{Parcels[Sites]} + \text{Plots[Sites}\wedge\text{Parcels]}$$

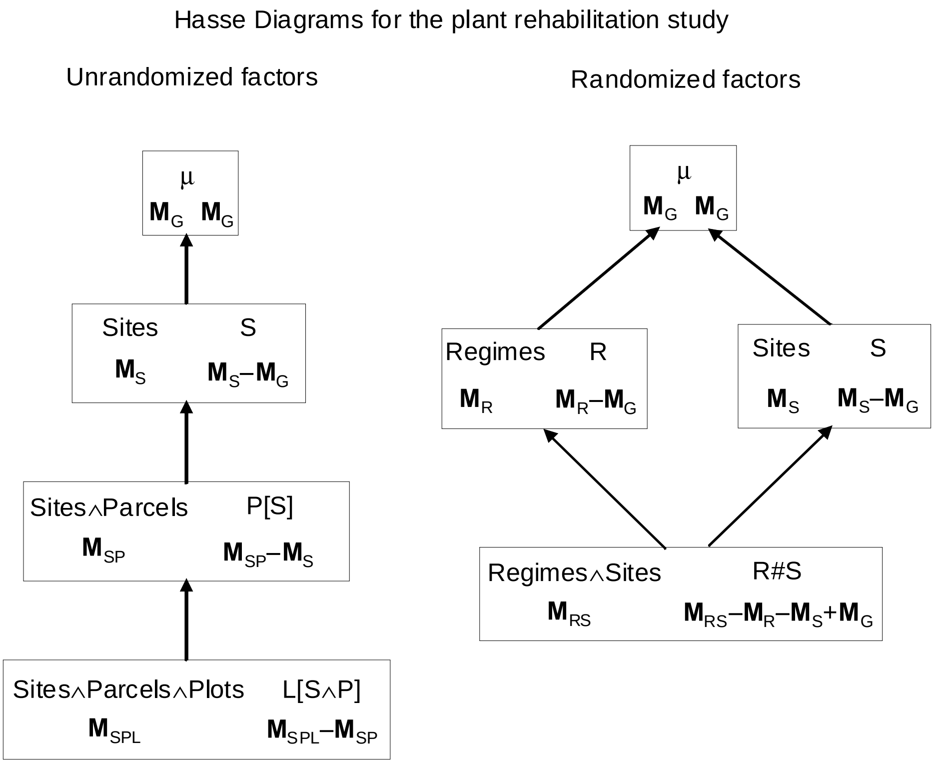
$$\text{Regimes*Sites} = \text{Regimes} + \text{Sites} + \text{Regimes\#Sites}$$

d) Degrees of freedom and sums of squares

The Hasse diagrams, with degrees of freedom, for this study are (use L for Plots to distinguish from P for Parcels):



The Hasse diagrams, with **M** and **Q** matrices, for this study are:



e) The analysis of variance table

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

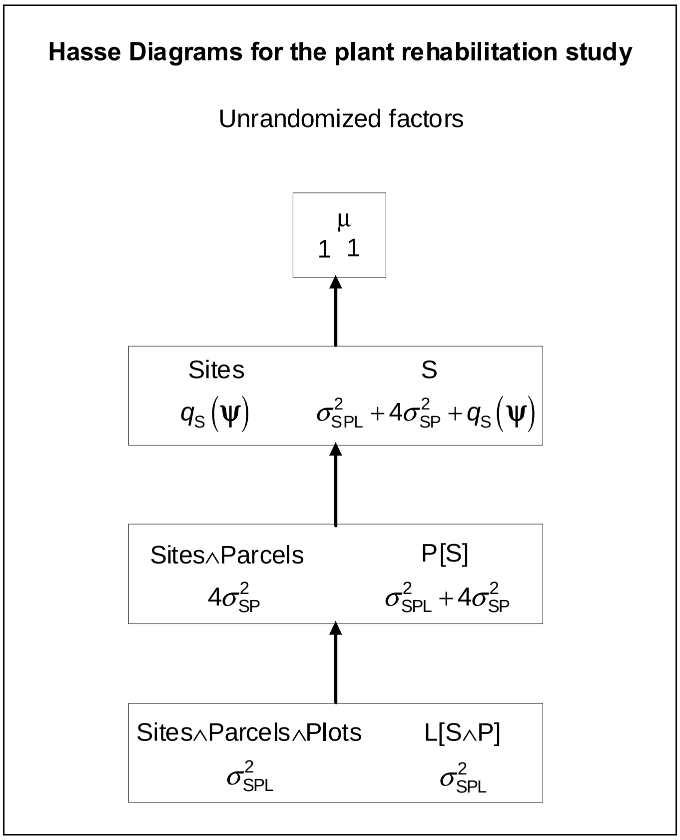
f) Maximal expectation and variation models

Seems that Sites and Regimes should be fixed and Parcels and Plots should be random. Hence, the maximal variation and expectation models are:

$$\mathrm{Var}[\mathrm{Y}] = \mathrm{Sites}\wedge\mathrm{Parcels} + \mathrm{Sites}\wedge\mathrm{Parcels}\wedge\mathrm{Plots}$$
$$\psi = \mathrm{E}[\mathrm{Y}] = \mathrm{Regimes}\wedge\mathrm{Sites}$$

g) The expected mean squares.

The Hasse diagram, with contributions to expected mean squares, for the unrandomized factors in this study are:



Note that all the contributions of the randomized generalized factors will be of the form $q_F(\boldsymbol{\psi})$ as all the randomized factors are fixed.

The analysis of variance table is:

Source	df	SSq	E[MSq]		
Sites	2	$\mathbf{Y}'\mathbf{Q}_{\mathrm{S}}\mathbf{Y}$	σ^2_{SPL}	$+4\sigma^2_{\mathrm{SP}}$	$+q_{\mathrm{S}}(\boldsymbol{\psi})$
Parcels[Sites]	12	$\mathbf{Y}'\mathbf{Q}_{\mathrm{SP}}\mathbf{Y}$	σ^2_{SPL}	$+4\sigma^2_{SP}$	
Plots[Sites∧Parcels]	45	$\mathbf{Y}'\mathbf{Q}_{\mathrm{SPL}}\mathbf{Y}$			
Regimes	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{R}}\mathbf{Y}$	σ^2_{SPL}		$+q_{\mathrm{R}}(\boldsymbol{\psi})$
Regimes#Sites	6	$\mathbf{Y}'\mathbf{Q}_{\mathrm{RS}}\mathbf{Y}$	σ^2_{SPL}		$+q_{\mathrm{RS}}(\boldsymbol{\psi})$
Residual	36	$\mathbf{Y}'\mathbf{Q}_{\mathrm{SPL}_{\mathrm{Res}}}\mathbf{Y}$	σ^2_{SPL}		

■

Example VI.13 Pollution effects of petrol additives

Suppose a study is to be conducted to investigate whether four petrol additives differ in the amount by which they reduce the emission of oxides of nitrogen. Four cars and four drivers are to be employed in the study and the following Latin square arrangement is to be used to assign the additives to the driver-car combinations:

		Car			
		1	2	3	4
	I	B	D	C	A
	II	A	B	D	C
Drivers	III	D	C	A	B
	IV	C	A	B	D

(Additives A, B, C, D)

a) *Description of pertinent features of the study*

1. the observational unit – a driver in a car
2. response variable – Reduction
3. unrandomized factors – Drivers, Cars
4. randomized factors – Additives
5. type of study – a Latin square

b) *The experimental structure*

Structure	Formula
unrandomized	*4* Drivers**4* Cars
randomized	*4* Additives

c) *Sources derived from the structure formulae*

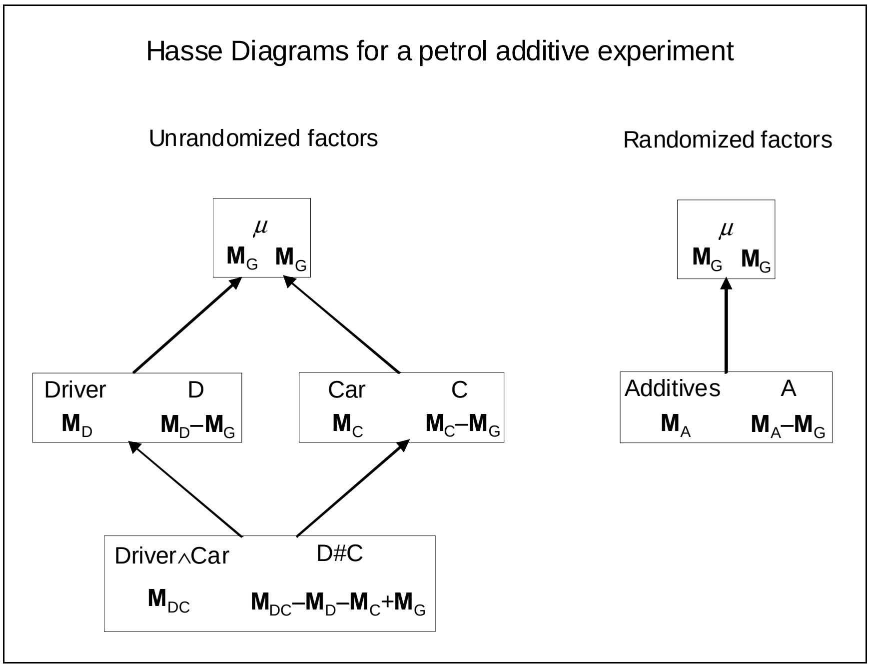
Drivers*Cars = Drivers + Cars + Drivers#Cars

Additives = Additives

d) Degrees of freedom and sums of squares

The degrees of freedom for this study can be worked out using the rule for completely crossed structures.

The Hasse diagrams, with **M** and **Q** matrices, for this study are:



e) The analysis of variance table

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

f) Maximal expectation and variation models

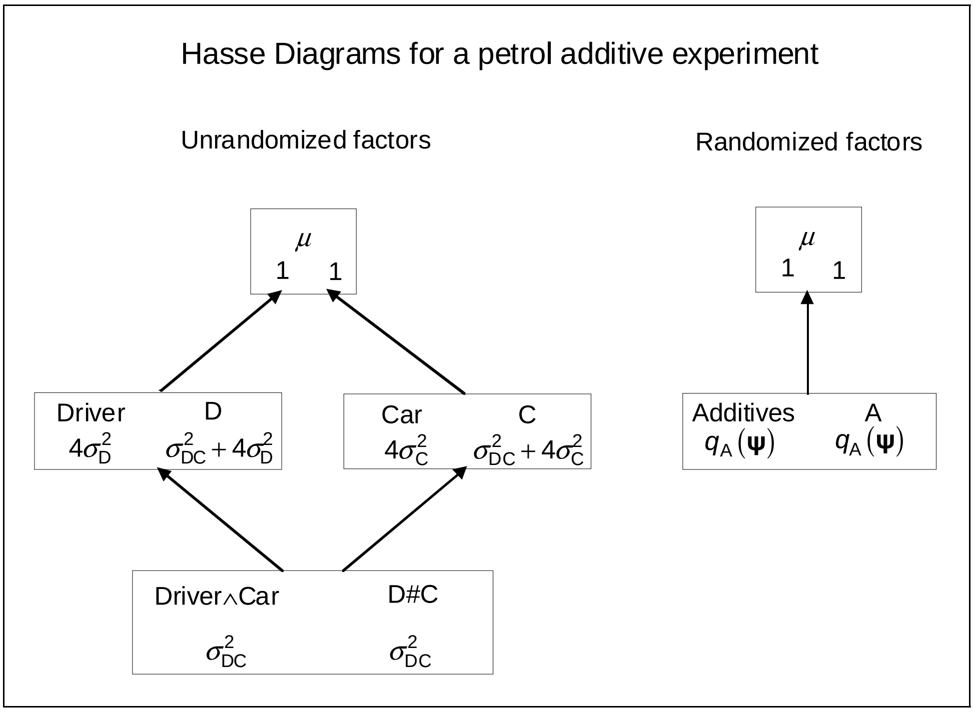
Will take it that randomized factors should be fixed and unrandomized factors should be random. Hence, the maximal variation and expectation models are:

$$\text{Var}[\mathbf{Y}] = \text{Drivers} + \text{Cars} + \text{Drivers}\wedge\text{Cars}$$

$$\psi = E[\mathbf{Y}] = \text{Additives}$$

g) The expected mean squares.

The Hasse diagrams, with contributions to expected mean squares, for this study are:



The analysis of variance table is:

Source	Df	SSq	E[MSq]
Driver	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{D}}\mathbf{Y}$	$\sigma^2_{DC} + 4\sigma^2_D$
Car	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{C}}\mathbf{Y}$	$\sigma^2_{DC} + 4\sigma^2_C$
Driver#Car	9	$\mathbf{Y}'\mathbf{Q}_{\mathrm{DC}}\mathbf{Y}$	
Additive	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{A}}\mathbf{Y}$	$\sigma^2_{\mathrm{DC}} + q_{\mathrm{A}}(\boldsymbol{\psi})$
Residual	6	$\mathbf{Y}'\mathbf{Q}_{\mathrm{DC}_{\mathrm{Res}}}\mathbf{Y}$	σ^2_{DC}
Total	15		

Suppose that the experiment is to be repeated by replicating the Latin square twice using the same cars but new drivers on a second occasion. What are the features of the study?

a) Description of pertinent features of the study

1. the observational unit – a driver in a car on an occasion
2. response variable – Reduction
3. unrandomized factors – Occasions, Drivers, Cars
4. randomized factors – Additives
5. type of study – Sets of Latin squares

b) The experimental structure

Structure	Formula
unrandomized	(*2* Occasions/*4* Drivers)**4* Cars
randomized	*4* Additives

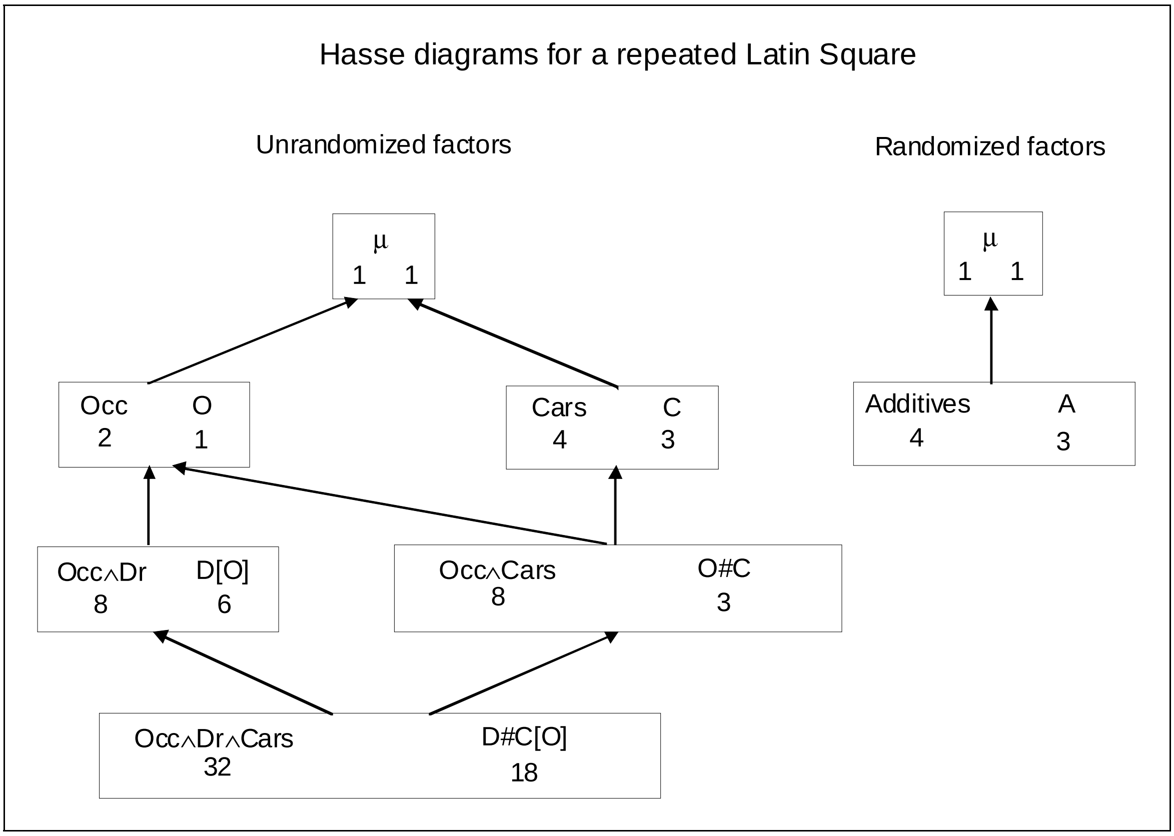
c) Sources derived from the structure formulae

(Occasions/Drivers)*Cars = Occasions + Drivers[Occasions] + Cars
+ Occasions #Cars + Drivers#Cars[Occasions]

Additives = Additives

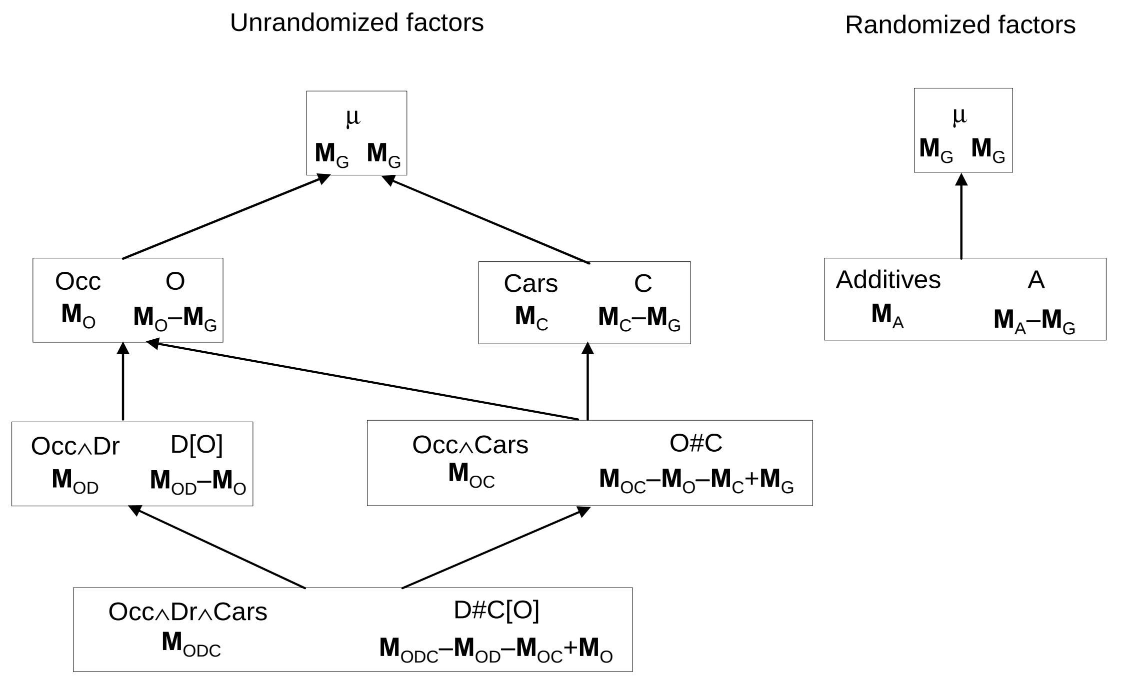
d) Degrees of freedom and sums of squares

The Hasse diagrams, with degrees of freedom, for this study are:



The Hasse diagrams, with **M** and **Q** matrices, for this study are:

Hasse diagrams for a repeated Latin Square



e) *The analysis of variance table*

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

f) *Maximal expectation and variation models*

Again, will take it that randomized factors should be fixed and unrandomized factors should be random. Hence, the maximal variation and expectation models are:

$$\text{Var}[\mathbf{Y}] = \text{Occasions} + \text{Occasions}\wedge\text{Drivers} + \text{Cars} + \text{Occasions}\wedge\text{Cars} + \text{Occasions}\wedge\text{Drivers}\wedge\text{Cars}$$

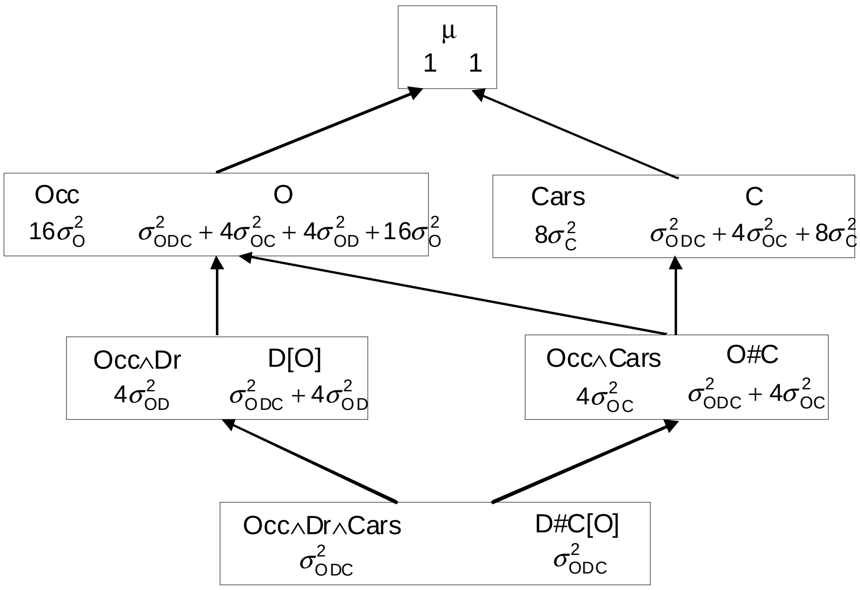
$$\psi = \text{E}[\mathbf{Y}] = \text{Additives}$$

g) *The expected mean squares.*

The Hasse diagram, with contributions to expected mean squares, for the unrandomized factors in this study are:

Hasse diagrams for a repeated Latin Square

Unrandomized factors



The analysis of variance table is:

Source	df	SSq	E[MSq]			
Occasions	1	$\mathbf{Y}'\mathbf{Q}_{\mathrm{O}}\mathbf{Y}$	σ^2_{ODC}	$+4\sigma^2_{\mathrm{OC}}$	$+4\sigma^2_{\mathrm{OD}}$	$+16\sigma^2_{\mathrm{O}}$
Dr[Occ]	6	$\mathbf{Y}'\mathbf{Q}_{\mathrm{OD}}\mathbf{Y}$	σ^2_{ODC}		$+4\sigma^2_{\mathrm{OD}}$	
Cars	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{C}}\mathbf{Y}$	σ^2_{ODC}	$+4\sigma^2_{\mathrm{OC}}$		$+8\sigma^2_{\mathrm{C}}$
Occ#Cars	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{OC}}\mathbf{Y}$	σ^2_{ODC}	$+4\sigma^2_{\mathrm{OC}}$		
Dr#Cars[Occ]	18	$\mathbf{Y}'\mathbf{Q}_{\mathrm{ODC}}\mathbf{Y}$				
Additives	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{A}}\mathbf{Y}$	σ^2_{ODC}			$+q_{\mathrm{A}}(\boldsymbol{\psi})$
Residual	15	$\mathbf{Y}'\mathbf{Q}_{\mathrm{ODC}_{\mathrm{Res}}}\mathbf{Y}$	σ^2_{ODC}			

■

Example VI.14 Controlled burning

Suppose an environmental scientist wants to investigate the effect on the biomass of burning Areas of natural vegetation. There are available two Areas separated by several kilometres for use in the investigation. It is only possible to either burn or not burn an entire Area. The scientist randomly selects to burn one Area and the other Area is left unburnt as a control. She randomly samples 30 locations in each Area and measures the biomass at each location.

a) Description of pertinent features of the study

3. the observational unit – a location
4. response variable – Biomass
5. unrandomized factors – Areas, Locations
6. randomized factors – Burning
7. type of study – a CRD with subsampling

b) The experimental structure

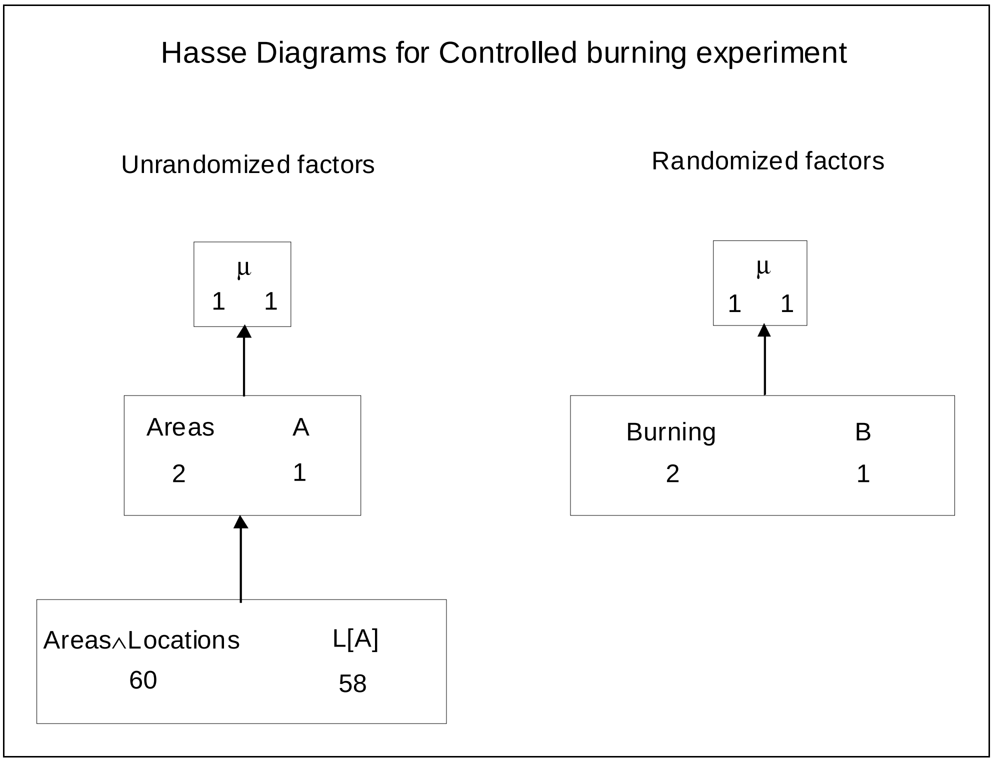
Structure	Formula
unrandomized	*2* Areas/*30* Locations
randomized	*2* Burning

c) Sources derived from the structure formulae

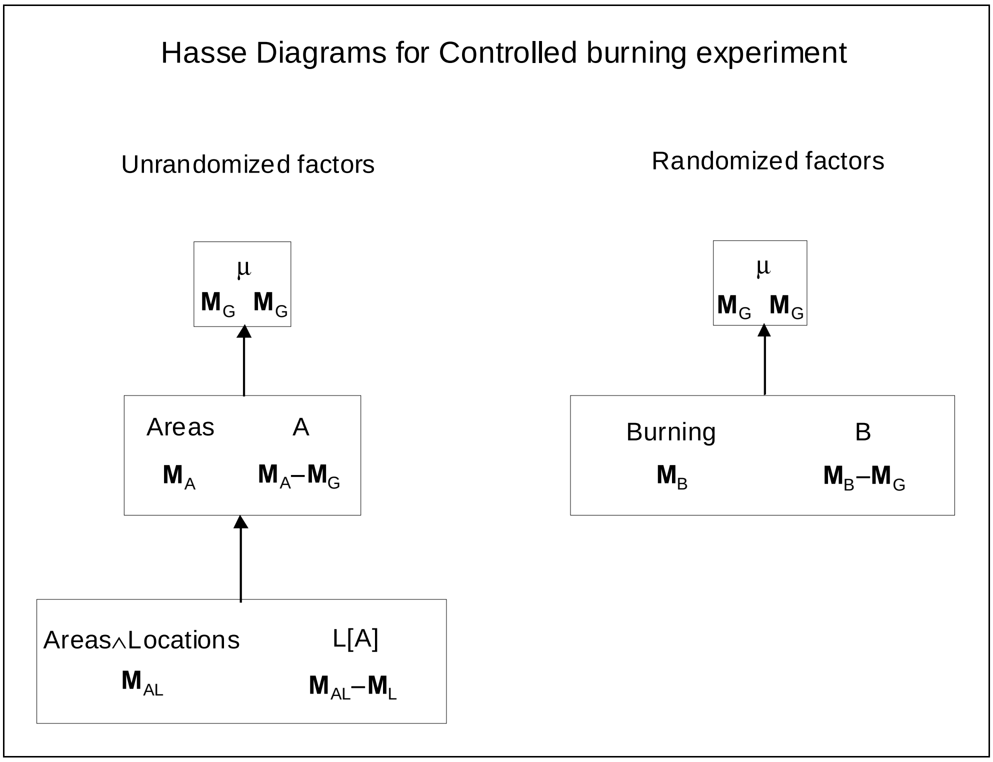
Areas/Locations = Areas + Locations[Areas]
Burning = Burning

d) Degrees of freedom and sums of squares

The Hasse diagrams, with degrees of freedom, for this study are:



The Hasse diagrams, with **M** and **Q** matrices, for this study are:



e) The analysis of variance table

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

f) Maximal expectation and variation models

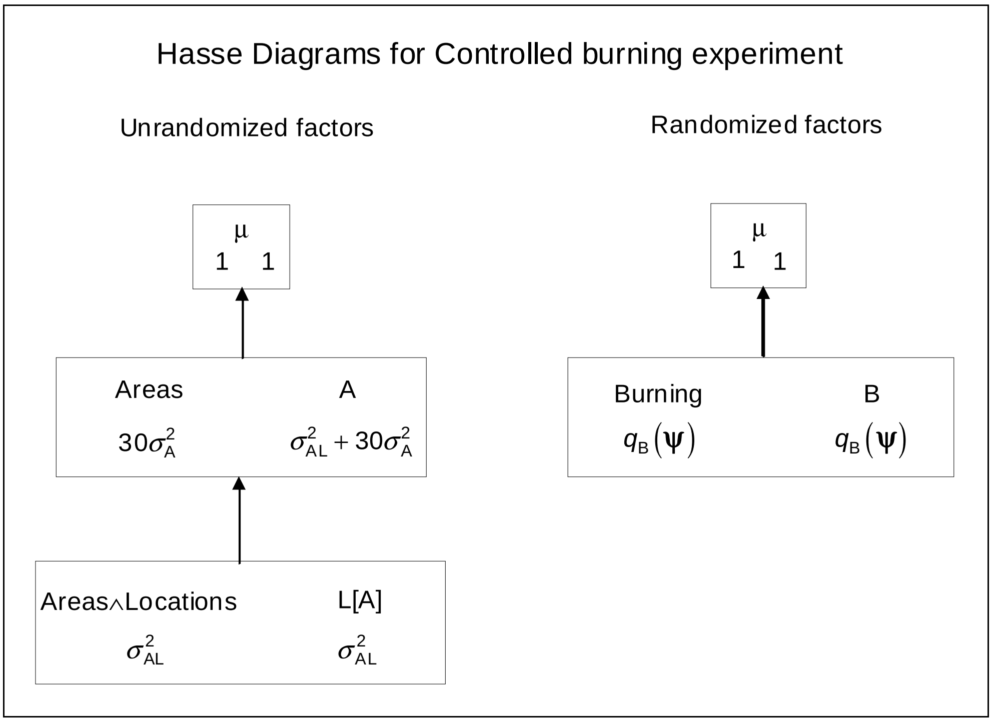
Seems that randomized factor should be fixed and unrandomized factors should be random. Hence, the maximal variation and expectation models are:

$$\text{Var}[\text{Y}] = \text{Areas} + \text{Areas}\wedge\text{Locations}$$

$$\psi = \text{E}[\text{Y}] = \text{Burning}$$

g) The expected mean squares.

The Hasse diagrams, with contributions to expected mean squares, for this study are:



The analysis of variance table is:

Source	df	SSq	E[MSq]
Areas	1	$\mathbf{Y}'\mathbf{Q}_{\text{A}}\mathbf{Y}$	
Burning	1	$\mathbf{Y}'\mathbf{Q}_{\text{B}}\mathbf{Y}$	$\sigma^2_{\text{AL}} + 30\sigma^2_{\text{A}} + q_{\text{B}}(\boldsymbol{\psi})$
Locations[Areas]	58	$\mathbf{Y}'\mathbf{Q}_{\text{AL}}\mathbf{Y}$	σ^2_{AL}

We see that we cannot test for Burning differences because there is no source with expected mean square $\sigma^2_{\text{AL}} + 30\sigma^2_{\text{A}}$. The problem is that Burning and Area differences are totally confounded and cannot be separated. This shows up in the analysis of variance table derived using our approach.

■

Example VI.15 Generalized randomized complete block design

A generalized randomized complete block design is the same as the ordinary randomized complete block design, except that each treatment occurs more than once in a block — see section IV.G, *Generalized randomized complete block design*. For example, suppose four treatments are to be compared when applied to a new variety of wheat. I employed a generalized randomized complete block design with 12 plots in each of 2 blocks so that each treatment is replicated 3 times in each block. The yield of wheat from each plot was measured. A possible layout for this experiment is shown in the table given below.

Layout for a generalized randomized complete block experiment

		Plots											
		1	2	3	4	5	6	7	8	9	10	11	12
Blocks	I	C	B	B	A	C	D	A	A	D	B	D	C
	II	C	B	A	D	D	D	A	A	B	C	B	C

Here work out the analysis for this experiment that includes a source for Block#Treatment interaction, and assumes that the unrandomized factors are random and the randomized factors are fixed. Having done this, derive the analysis for only Plots random and the rest of the factors fixed.

a) *Description of pertinent features of the study*

1. the observational unit – a plot
2. response variable – Yield
3. unrandomized factors – Blocks, Plots
4. randomized factors – Treatments
5. type of study – Generalized Randomized Complete Block Design

b) *The experimental structure*

Structure	Formula
unrandomized	*2* Blocks/*12* Plots
randomized	*4* Treatments*Blocks

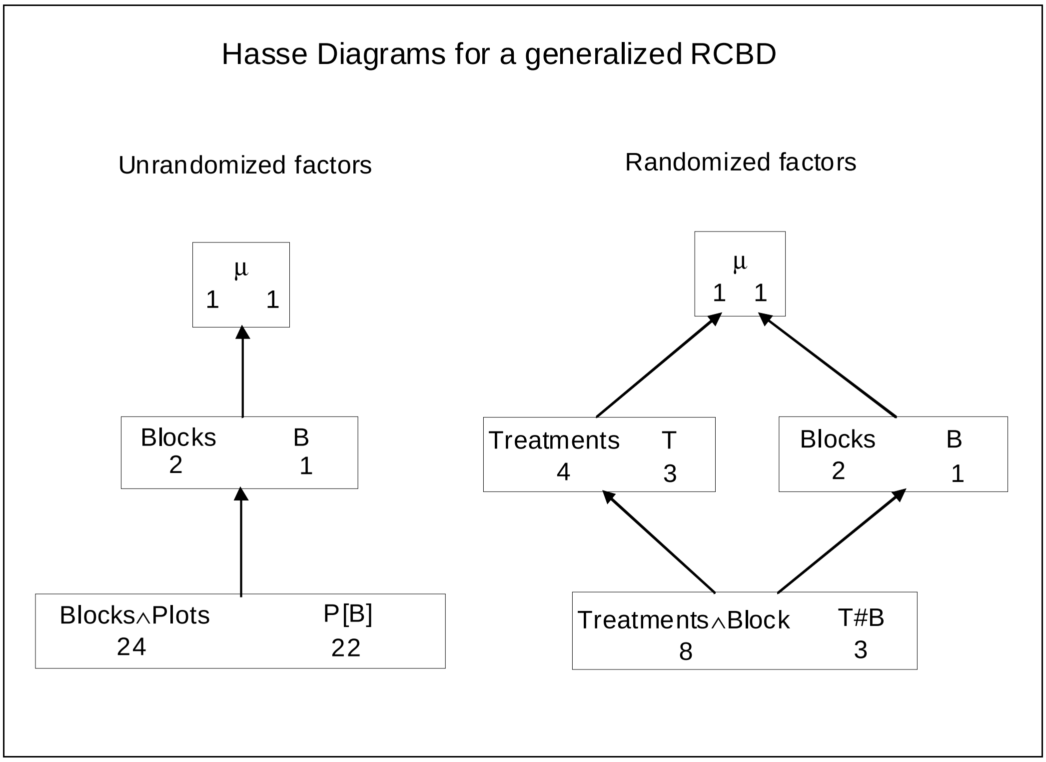
c) *Sources derived from the structure formulae*

Blocks/Plots = Blocks + Plots[Blocks]

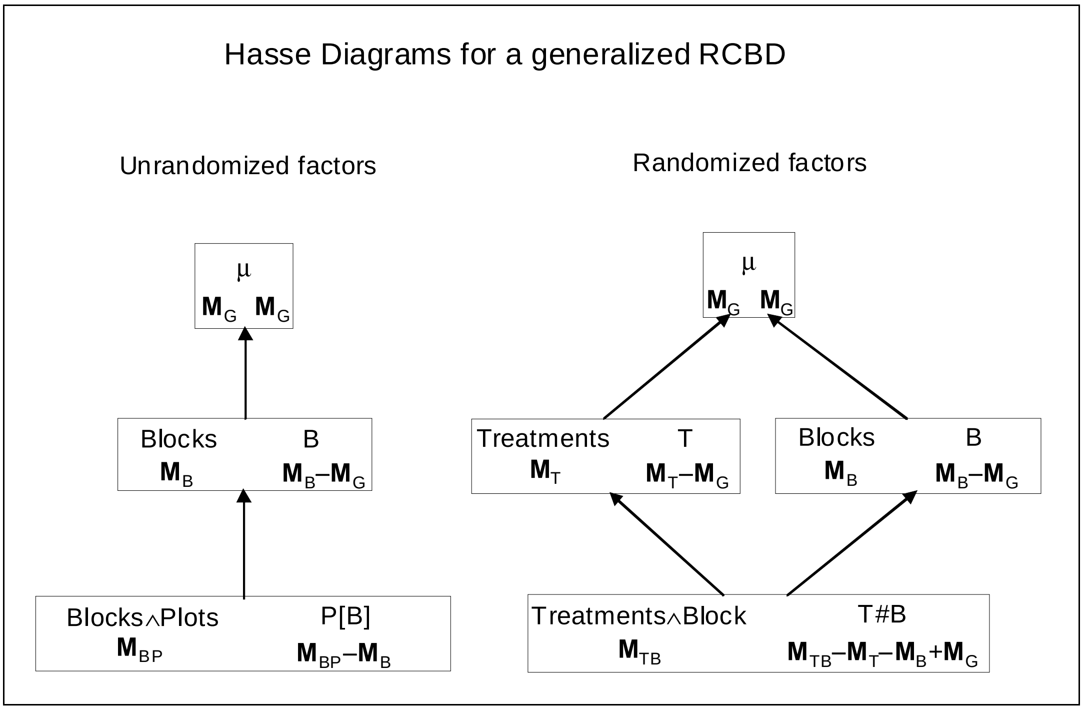
Treatments*Blocks = Treatments + Blocks + Treatments#Blocks

d) *Degrees of freedom and sums of squares*

The Hasse diagrams, with degrees of freedom, for this study are:



The Hasse diagrams, with **M** and **Q** matrices, for this study are:



e) The analysis of variance table

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

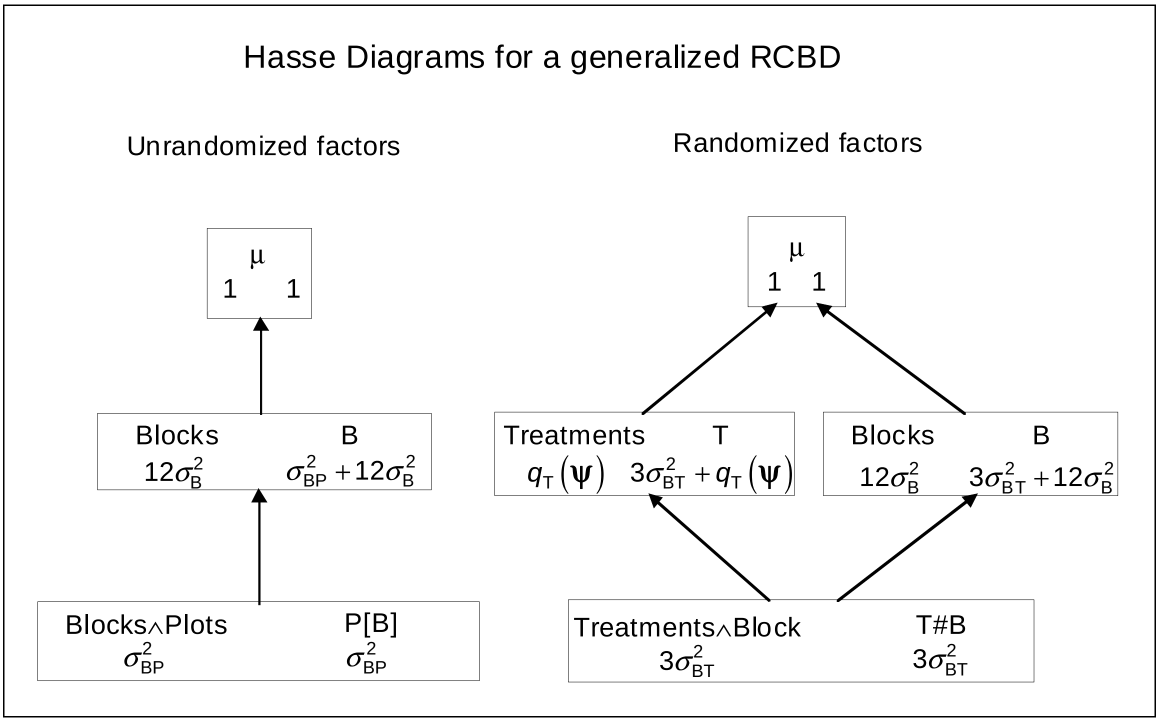
f) *Maximal expectation and variation models*

Take the randomized factors to be fixed and unrandomized factors to be random. Hence, the maximal variation and expectation models are:

$$\mathrm{Var}[\mathrm{Y}] = \text{Blocks} + \text{Blocks}\wedge\text{Plots} + \text{Treatments}\wedge\text{Blocks}$$
$$\psi = \mathrm{E}[\mathrm{Y}] = \text{Treatments}$$

g) *The expected mean squares.*

The Hasse diagrams, with contributions to expected mean squares, for this study are:



The analysis of variance table is:

Source	df	SSq	E[MSq]
Blocks	1	$\mathbf{Y}'\mathbf{Q}_{\mathrm{B}}\mathbf{Y}$	$\sigma^2_{\mathrm{BP}} + 3\sigma^2_{\mathrm{BT}} + 12\sigma^2_{\mathrm{B}}$
Plots[Blocks]	22	$\mathbf{Y}'\mathbf{Q}_{\mathrm{BP}}\mathbf{Y}$	
Treatments	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{T}}\mathbf{Y}$	$\sigma^2_{\mathrm{BP}} + 3\sigma^2_{\mathrm{BT}} + q_{\mathrm{T}}(\boldsymbol{\psi})$
Treatments#Blocks	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{TB}}\mathbf{Y}$	$\sigma^2_{\mathrm{BP}} + 3\sigma^2_{\mathrm{BT}}$
Residual	16	$\mathbf{Y}'\mathbf{Q}_{\mathrm{BP_{Res}}}\mathbf{Y}$	σ^2_{BP}
Total	23		

For only Plots random, the maximal variation and expectation models are:

$$\text{Var}[\mathbf{Y}] = \text{Blocks}\wedge\text{Plots}$$
$$\psi = \text{E}[\mathbf{Y}] = \text{Treatments}\wedge\text{Blocks}$$

In this case there will be the one variation component, σ^2_{BP}. The remaining contributions will be of the form $q_{\text{F}}(\boldsymbol{\psi})$.

The analysis of variance table is:

Source	df	SSq	E[MSq]
Blocks	1	$\mathbf{Y}'\mathbf{Q}_{\text{B}}\mathbf{Y}$	$\sigma^2_{\text{BP}} + q_{\text{B}}(\boldsymbol{\psi})$
Plots[Blocks]	22	$\mathbf{Y}'\mathbf{Q}_{\text{BP}}\mathbf{Y}$	
Treatments	3	$\mathbf{Y}'\mathbf{Q}_{\text{T}}\mathbf{Y}$	$\sigma^2_{\text{BP}} + q_{\text{T}}(\boldsymbol{\psi})$
Treatments#Blocks	3	$\mathbf{Y}'\mathbf{Q}_{\text{TB}}\mathbf{Y}$	$\sigma^2_{\text{BP}} + q_{\text{BT}}(\boldsymbol{\psi})$
Residual	16	$\mathbf{Y}'\mathbf{Q}_{\text{BP}_{\text{Res}}}\mathbf{Y}$	σ^2_{BP}
Total	23		

Note the difference in the denominator of the test for Treatments between the two analyses. For the former analysis it is Treatments#Blocks and the latter analysis it is the Residual. Which analysis is correct depends on the nature of the Block-Treatment interaction. If the Blocks are very different, as in the case where they are quite different sites, and it is anticipated that the treatments will respond quite differently in the different blocks, then it is probable that the Blocks should be regarded as fixed so that the term Treatments#Blocks is also fixed (see Steel and Torrie, sec. 9.8).

■

Example VI.16 Salt tolerance of lizards

To examine the salt tolerance of the lizard *Tiliqua rugosa*, eighteen lizards of this species were obtained. Each lizard was randomly selected to receive one of three salt treatments (injection with sodium, injection with potassium, no injection) so that 6 lizards received each treatment. Blood samples were then taken from each lizard on five occasions after injection and the concentration of Na in the sample determined.

a) Description of pertinent features of the study

1. the observational unit – a lizard on an occasion
2. response variable – Na concentration
3. unrandomized factors – Lizards, Occasions
4. randomized factors – Treatments
5. type of study – a repeated measures CRD

b) *The experimental structure*

Structure	Formula
unrandomized	*18* Lizards*5 Occasions
randomized	*3* Treatments*Occasions

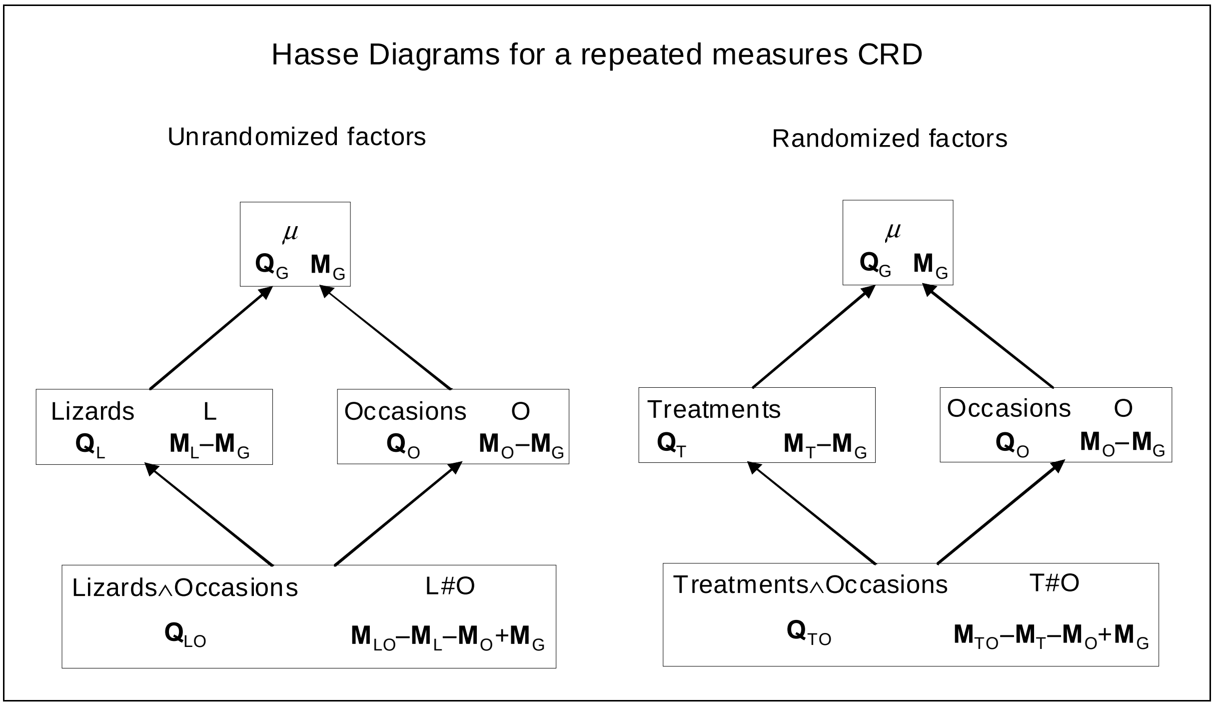
c) *Sources derived from the structure formulae*

Lizards*Occasions = Lizards + Occasions + Lizards#Occasions

Treatments*Occasions = Treatments + Occasions + Treatments#Occasions

d) *Degrees of freedom and sums of squares*

The degrees of freedom for this study can be worked out using the rule for completely crossed structures. The Hasse diagrams, with **M** and **Q** matrices, for this study are:



e) *The analysis of variance table*

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

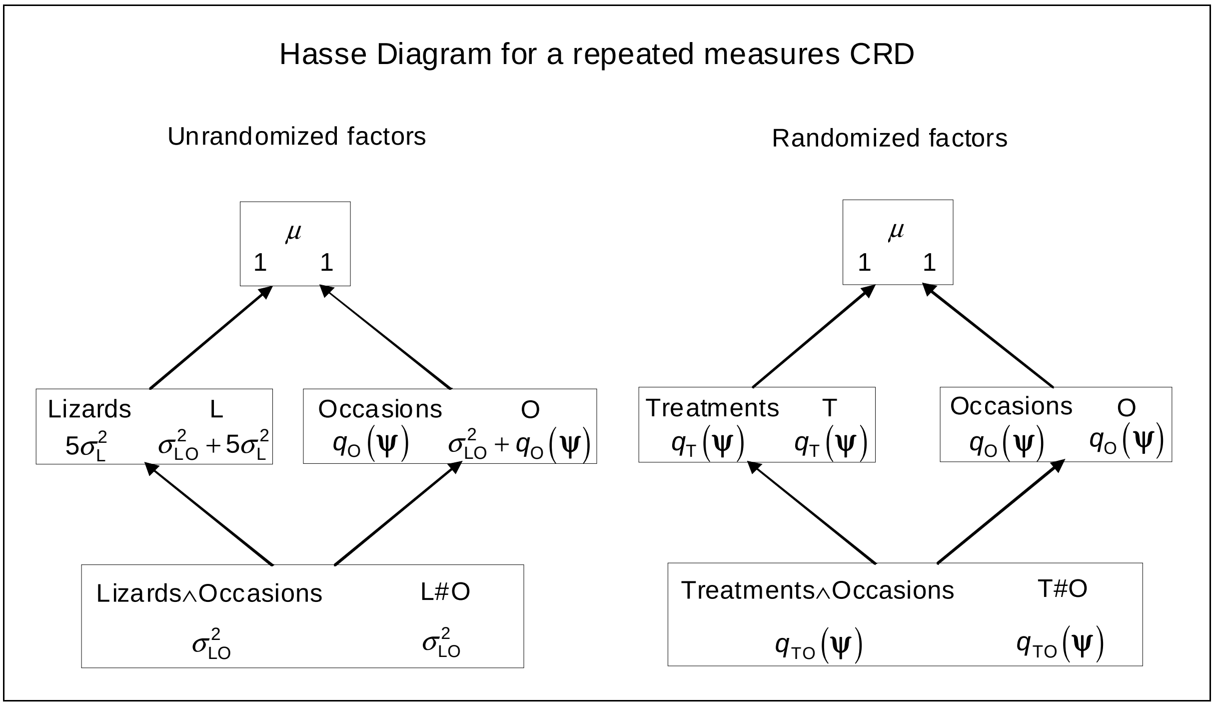
f) *Maximal expectation and variation models*

Seems that Treatments and Occasions should be fixed and Lizards should be random. Hence, the maximal variation and expectation models are:

$$\text{Var}[Y] = \text{Lizards} + \text{Lizards}\wedge\text{Occasions}$$
$$\psi = \text{E}[Y] = \text{Treatments}\wedge\text{Occasions}$$

g) The expected mean squares.

The Hasse diagrams, with contributions to expected mean squares, for this study are:



The analysis of variance table is:

Source	df	SSq	E[MSq]		
Occasions	4	$\mathbf{Y}'\mathbf{Q}_{\mathrm{O}}\mathbf{Y}$	σ^2_{LO}		$+q_{\mathrm{O}}(\boldsymbol{\psi})$
Lizards	17	$\mathbf{Y}'\mathbf{Q}_{\mathrm{L}}\mathbf{Y}$			
Treatments	2	$\mathbf{Y}'\mathbf{Q}_{\mathrm{T}}\mathbf{Y}$	σ^2_{LO}	$+5\sigma^2_{\mathrm{L}}$	$+q_{\mathrm{T}}(\boldsymbol{\psi})$
Residual	15	$\mathbf{Y}'\mathbf{Q}_{\mathrm{L_{Res}}}\mathbf{Y}$	σ^2_{LO}	$+5\sigma^2_{\mathrm{L}}$	
Lizards#Occasions	68	$\mathbf{Y}'\mathbf{Q}_{\mathrm{LO}}\mathbf{Y}$			
Treatments#Occasions	8	$\mathbf{Y}'\mathbf{Q}_{\mathrm{TO}}\mathbf{Y}$	σ^2_{LO}		$+q_{\mathrm{TO}}(\boldsymbol{\psi})$
Residual	60	$\mathbf{Y}'\mathbf{Q}_{\mathrm{LO_{Res}}}\mathbf{Y}$	σ^2_{LO}		

■

Example VI.17 Eucalyptus growth

An experiment was planted in a forest in Queensland to study the effects of irrigation and fertilizer on 4 seedlots of a species of gum tree. There were two levels of irrigation (no and yes), two levels of fertilizer (no and yes) and four seedlots (Bulahdelah, Coffs Harbour, Pomona and Atherton). Because of the difficulties of irrigating and applying fertilizers to individual trees, these needed to be applied to groups of trees. So the experimental area was divided up into 8 stands of 20 trees, with four stands in one block and the other four in a second block. The four combinations of irrigation and fertilizer were randomized to the four stands in a block. Each stand of 20 trees consisted of 4 rows by 5 columns and the 4 seedlots were randomized to the four rows. The mean height of the five trees in a row was measured.

a) *Description of pertinent features of the study*

1. the observational unit – a row of trees
2. response variable – Mean height
3. unrandomized factors – Blocks, Stands, Rows
4. randomized factors – Irrigation, Fertilizer, Seedlots
5. type of study – a split-plot design with main plots in a RCBD
6. and subplots completely randomized

b) *The experimental structure*

Structure	Formula
unrandomized	*2* Blocks/*4* Stands/*4* Rows
randomized	*2* Irrigation**2* Fertilizer**4* Seedlots

c) *Sources derived from the structure formulae*

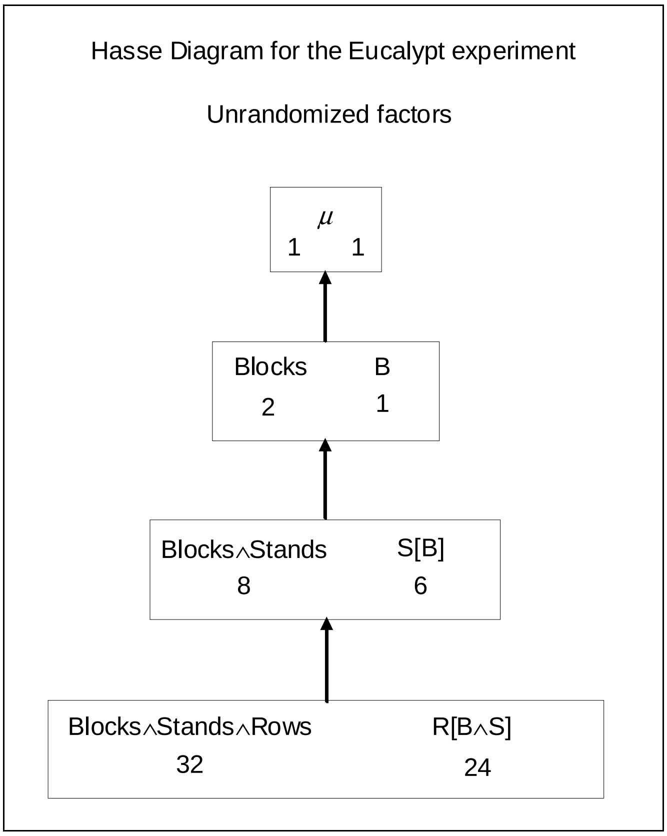
Blocks/Stands/Rows = Blocks + Stands[Blocks] + Rows[Blocks$\wedge$Stands]

Irrigation*Fertilizer*Seedlots = (Irrigation + Fertilizer + Irrigation#Fertilizer)*Seedlot
= Irrigation + Fertilizer + Irrigation#Fertilizer
+ Irrigation#Seedlots + Fertilizer#Seedlots
+ Irrigation#Fertilizer#Seedlots

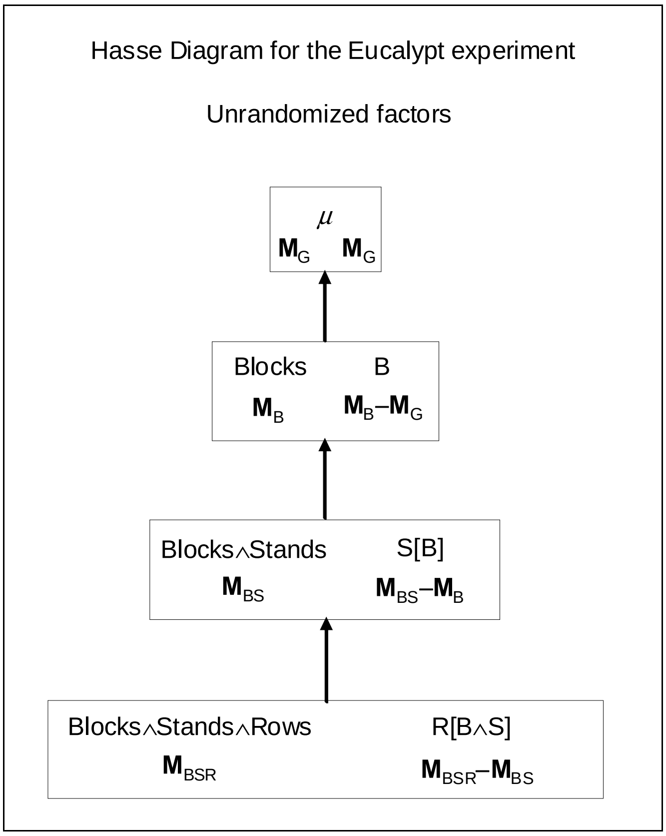
d) *Degrees of freedom and sums of squares*

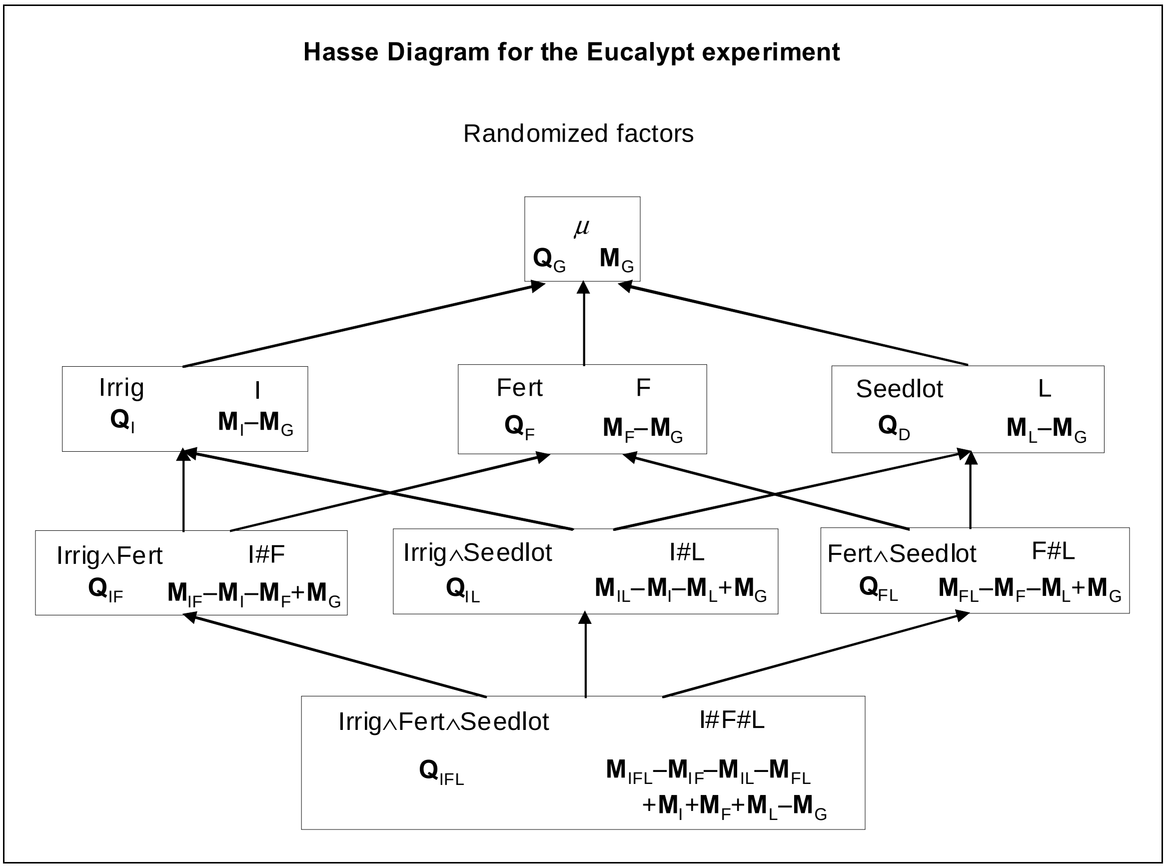
The Hasse diagrams, with degrees of freedom, for this study are:

In this case we only do the Hasse diagram for the unrandomized factors because the degrees of freedom for the randomized factors can be obtained using the rule for all factors crossed.



The Hasse diagrams, with **M** and **Q** matrices, for this study are:





Note the use of "L" for Seedlot to distinguish it from "S" for Stand.

e) The analysis of variance table

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

f) Maximal expectation and variation models

Take all the randomized factors and Blocks to be fixed; the remainder of the factors take as random. Hence, the maximal variation and expectation models are:

$$\mathrm{Var}[Y] = \mathrm{Blocks}\wedge\mathrm{Stands} + \mathrm{Blocks}\wedge\mathrm{Stands}\wedge\mathrm{Rows}$$

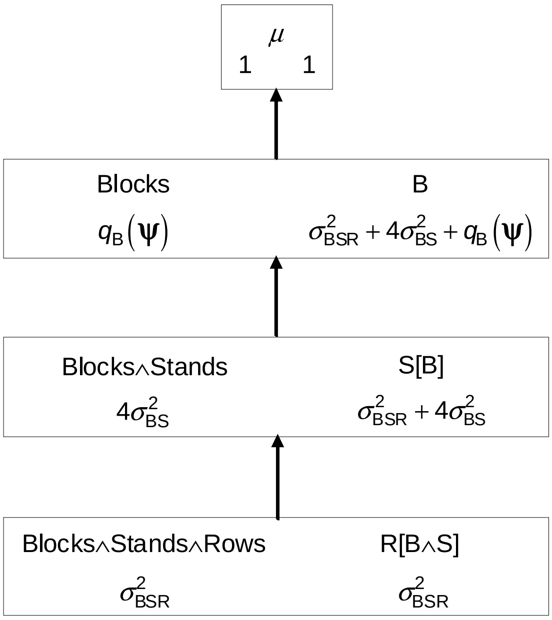
$$\psi = \mathrm{E}[Y] = \mathrm{Blocks} + \mathrm{Irrigation}\wedge\mathrm{Fertilizer}\wedge\mathrm{Seedlots}$$

g) The expected mean squares.

The Hasse diagram, with contributions to expected mean squares, for the unrandomized factors in this study are:

Hasse Diagram for the Eucalypt experiment

Unrandomized factors



Note that the contributions of all the randomized generalized factors will be of the form $q_F(\boldsymbol{\psi})$ as all the randomized factors are fixed.

The analysis of variance table is:

Source	df	SSq	E[MSq]
Blocks	1	$\mathbf{Y}'\mathbf{Q}_{\mathrm{B}}\mathbf{Y}$	$\sigma^2_{\mathrm{BSR}} + 4\sigma^2_{\mathrm{BS}} + q_{\mathrm{B}}(\boldsymbol{\psi})$
Stands[Blocks]	6	$\mathbf{Y}'\mathbf{Q}_{\mathrm{BS}}\mathbf{Y}$	
Irrigation	1	$\mathbf{Y}'\mathbf{Q}_{\mathrm{I}}\mathbf{Y}$	$\sigma^2_{\mathrm{BSR}} + 4\sigma^2_{\mathrm{BS}} + q_{\mathrm{I}}(\boldsymbol{\psi})$
Fertilizer	1	$\mathbf{Y}'\mathbf{Q}_{\mathrm{F}}\mathbf{Y}$	$\sigma^2_{\mathrm{BSR}} + 4\sigma^2_{\mathrm{BS}} + q_{\mathrm{F}}(\boldsymbol{\psi})$
Irrigation#Fertilizer	1	$\mathbf{Y}'\mathbf{Q}_{\mathrm{IF}}\mathbf{Y}$	$\sigma^2_{\mathrm{BSR}} + 4\sigma^2_{\mathrm{BS}} + q_{\mathrm{IF}}(\boldsymbol{\psi})$
Residual	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{BS}_{\mathrm{Res}}}\mathbf{Y}$	$\sigma^2_{\mathrm{BSR}} + 4\sigma^2_{\mathrm{BS}}$
Rows[Blocks∧Stands]	24	$\mathbf{Y}'\mathbf{Q}_{\mathrm{BSR}}\mathbf{Y}$	
Seedlots	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{L}}\mathbf{Y}$	$\sigma^2_{\mathrm{BSR}} + q_{\mathrm{L}}(\boldsymbol{\psi})$
Irrigation#Seedlots	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{IL}}\mathbf{Y}$	$\sigma^2_{\mathrm{BSR}} + q_{\mathrm{IL}}(\boldsymbol{\psi})$
Fertilizer#Seedlots	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{FL}}\mathbf{Y}$	$\sigma^2_{\mathrm{BSR}} + q_{\mathrm{FL}}(\boldsymbol{\psi})$
Irrigation#Fertilizer#Seedlot	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{IFL}}\mathbf{Y}$	$\sigma^2_{\mathrm{BSR}} + q_{\mathrm{IFL}}(\boldsymbol{\psi})$
Residual	12	$\mathbf{Y}'\mathbf{Q}_{\mathrm{BSR}_{\mathrm{Res}}}\mathbf{Y}$	σ^2_{BSR}

■

Example VI.18 Eelworm experiment (not examinable)

Cochran and Cox (1957, section 3.2) present the results of an experiment examining the effects of soil fumigants on the number of eelworms. There were four different fumigants each applied in both single and double dose rates as well as a control treatment in which no fumigant was applied. The experiment was laid out in 4 blocks each containing 12 plots; in each block, the 8 treatment combinations were each applied once and the control treatment four times and the 12 treatments randomly allocated to plots. The number of eelworm cysts in 400g samples of soil from each plot was determined.

a) *Description of pertinent features of the study*

1. the observational unit – a plot
2. response variable – Number of cysts
3. unrandomized factors – Blocks, Plots
4. randomized factors – Control, Fumigant, Dose
5. type of study – an RCBD

b) *The experimental structure*

Structure	Formula
unrandomized	*4* Blocks/*12* Plots
randomized	*2* Control/(*4* Fumigant**2* Dose)

c) *Sources derived from the structure formulae*

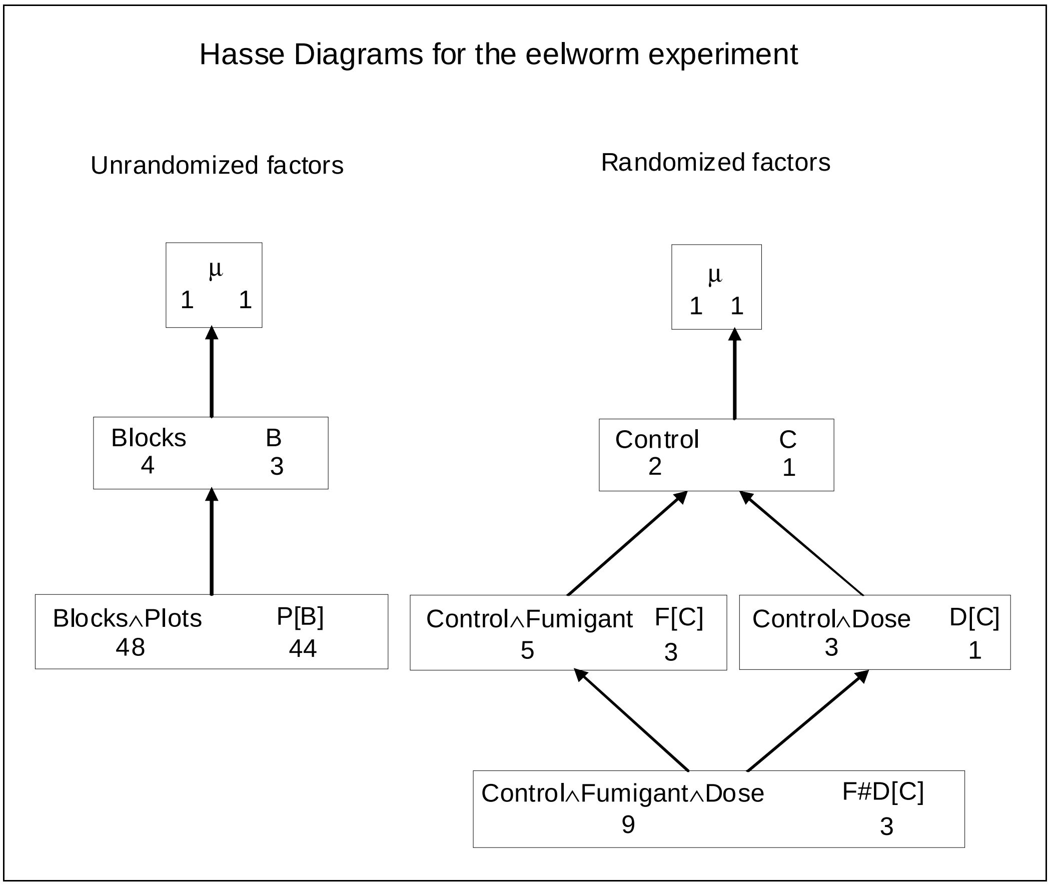
Blocks/Plots = Blocks + Plots[Blocks]

Control/(Fumigant*Dose)

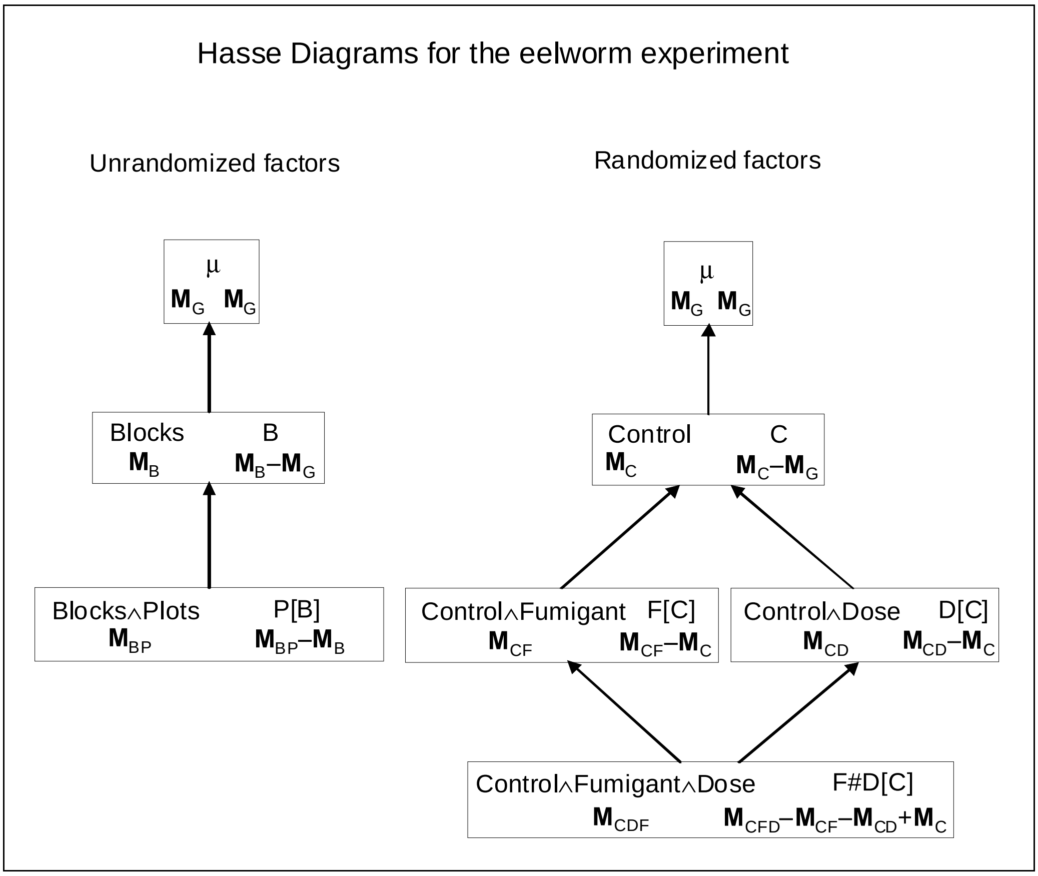
= Control/(Fumigant + Dose + Fumigant#Dose)
= Control + Fumigant[Control] + Dose[Control]
+ Fumigant#Dose[Control]

d) *Degrees of freedom and sums of squares*

The Hasse diagrams, with degrees of freedom, for this study are:



The Hasse diagrams, with **M** and **Q** matrices, for this study are:



e) The analysis of variance table

Enter the sources for the study, their degrees of freedom and quadratic forms, into the analysis of variance table below.

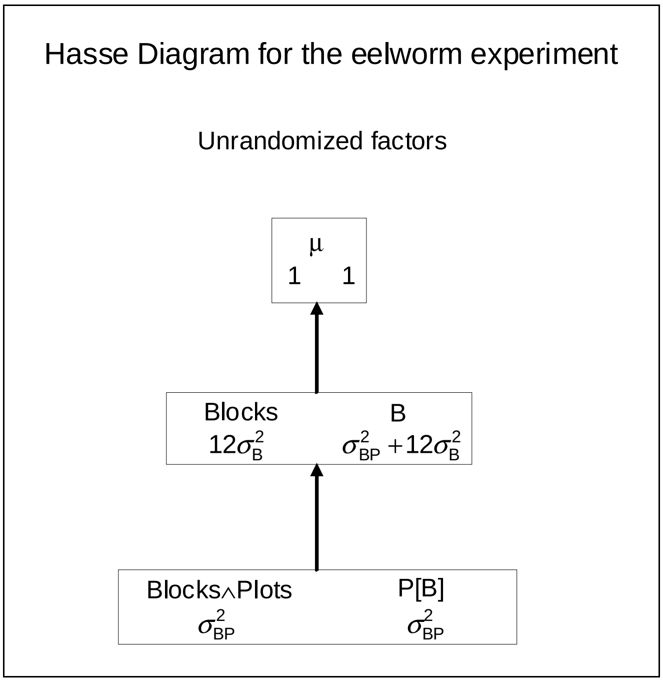
f) Maximal expectation and variation models

Take the randomized factors to be fixed and unrandomized factors to be random. Hence, the maximal variation and expectation models are:

$$\text{Var[Y]} = \text{Blocks} + \text{Blocks}\wedge\text{Plots}$$
$$\psi = \text{E[Y]} = \text{Control}\wedge\text{Dose}\wedge\text{Fumigant}$$

g) The expected mean squares.

The Hasse diagram, with contributions to expected mean squares, for the unrandomized factors in this study are:



Note that the contributions of all the randomized generalized factors will be of the form $q_F(\boldsymbol{\psi})$ as all the randomized factors are fixed.

The analysis of variance table is:

Source	df	SSq	E[MSq]
Blocks	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{B}}\mathbf{Y}$	$\sigma^2_{\mathrm{BP}} + 12\sigma^2_{\mathrm{B}}$
Plots[Blocks]	44	$\mathbf{Y}'\mathbf{Q}_{\mathrm{BP}}\mathbf{Y}$	
Control	1	$\mathbf{Y}'\mathbf{Q}_{\mathrm{C}}\mathbf{Y}$	$\sigma^2_{\mathrm{BP}} + q_{\mathrm{C}}(\boldsymbol{\psi})$
Dose[Control]	1	$\mathbf{Y}'\mathbf{Q}_{\mathrm{CD}}\mathbf{Y}$	$\sigma^2_{\mathrm{BP}} + q_{\mathrm{CD}}(\boldsymbol{\psi})$
Fumigant[Control]	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{CF}}\mathbf{Y}$	$\sigma^2_{\mathrm{BP}} + q_{\mathrm{CF}}(\boldsymbol{\psi})$
Dose#Fumigant[Control]	3	$\mathbf{Y}'\mathbf{Q}_{\mathrm{CDF}}\mathbf{Y}$	$\sigma^2_{\mathrm{BP}} + q_{\mathrm{CDF}}(\boldsymbol{\psi})$
Residual	36	$\mathbf{Y}'\mathbf{Q}_{\mathrm{BP_{Res}}}\mathbf{Y}$	σ^2_{BP}
Total	47		

■

Example VI.19 A factorial experiment

An experiment is to be conducted on sugar cane to investigate 6 factor (A, B, C, D, E, F) each at two levels. This experiment is to involve 16 blocks each of eight plots. The 64 treatment combinations are divided into 8 sets of 8 so that the ABCD, ABEF and ACE interactions are associated with set differences. The 8 sets are randomized to the 16 blocks so that each set occurs on two blocks and the 8 combinations in a set are randomized to the plots within a block. The sugar content of the cane is to be measured.

a) Description of pertinent features of the study

1. the observational unit – a plot
2. response variable – Sugar content
3. unrandomized factors – Blocks, Plots
4. randomized factors – A, B, C, D, E, F
5. type of study – two replicates of a confounded 2^k factorial

b) The experimental structure

Structure	Formula
unrandomized	*16* Blocks/*8* Plots
randomized	*2* A**2* B**2* C**2* D**2* E**2* F

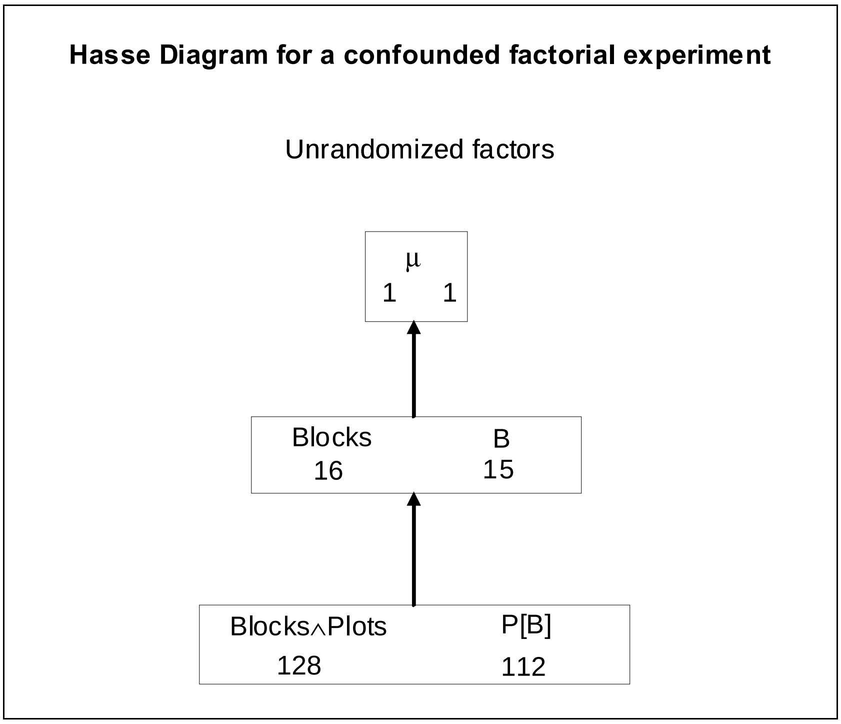
How many main effects, two factor interactions, three-factor interactions and interactions of more than 3 factors are there? What are the interactions confounded with blocks?

There are 6 main effects, 15 two-factor interactions, 20 three-factor interactions and 64 – 1 – 6 – 15 – 20 = 22 other interactions.

The interactions confounded with blocks are ABCD, ABEF and ACE and all the products of these. That is ABCD, ABEF and ACE and CDEF, BDE, BCF, ADF *c) Degrees of freedom and sums of squares*

What are the degrees of freedom for the unrandomized sources and for the randomized sources?

For the unrandomized generalized factors, the Hasse diagram is as follows:



For the randomized sources, by the crossed structure formula rule, all degrees of freedom will be 1 as all factors have 2 levels.

d) The analysis of variance table

Enter the sources for the study, and their degrees of freedom, into the analysis of variance table below.

e) Maximal expectation and variation models

Seems that randomized factors should be fixed and unrandomized factors should be random. Hence, the maximal variation and expectation models are:

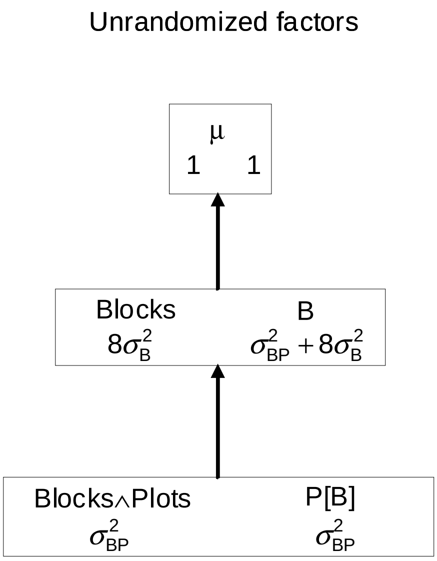
$$\mathrm{Var}[Y] = \mathrm{Blocks} + \mathrm{Blocks}\wedge\mathrm{Plots}$$

$$\psi = E[Y] = A\wedge B\wedge C\wedge D\wedge E\wedge F$$

f) The expected mean squares.

The Hasse diagram, with contributions to expected mean squares, for the unrandomized factors in this study are:

Hasse Diagram for a confounded factorial experiment



Note that the contributions of all the randomized generalized factors will be of the form $q_F(\boldsymbol{\psi})$ as all the randomized factors are fixed.

The analysis of variance table is (just give the interactions confounded with blocks and the numbers of main effects, two-factor, three-factor and other interactions):

Source	df	E[MSq]
Blocks	15	
ACE	1	$\sigma_{BP}^2 + 8\sigma_B^2 + q_{\mathbf{ACE}}(\psi)$
ADF	1	$\sigma_{BP}^2 + 8\sigma_B^2 + q_{\mathbf{ADF}}(\psi)$
BCF	1	$\sigma_{BP}^2 + 8\sigma_B^2 + q_{\mathbf{BCF}}(\psi)$
BDE	1	$\sigma_{BP}^2 + 8\sigma_B^2 + q_{\mathbf{BDE}}(\psi)$
ABCD	1	$\sigma_{BP}^2 + 8\sigma_B^2 + q_{\mathbf{ABCD}}(\psi)$
ABEF	1	$\sigma_{BP}^2 + 8\sigma_B^2 + q_{\mathbf{ABEF}}(\psi)$
CDEF	1	$\sigma_{BP}^2 + 8\sigma_B^2 + q_{\mathbf{CDEF}}(\psi)$
Residual	8	$\sigma_{BP}^2 + 8\sigma_B^2$
Plots[Blocks]	112	
main effects	6	$\sigma_{BP}^2 + q_{\mathbf{i}}(\psi)$
2-factor interactions	15	$\sigma_{BP}^2 + q_{\mathbf{i.j}}(\psi)$
3-factor interactions	17	$\sigma_{BP}^2 + q_{\mathbf{i.j.k}}(\psi)$
other interactions	19	$\sigma_{BP}^2 + q_{\mathbf{i.j.k.l+}}(\psi)$
Residual	55	σ_{BP}^2

■