Design and analysis of experiments testing for diversity in ecology

R. A. Bailey

Queen Mary University of London

r.a.bailey@qmul.ac.uk

Joint work with Julia Reiss, University of Roehampton

3rd International Conference on Design of Experiments, Memphis, USA, May 2011

Abstract

It is now widely believed that biological diversity is good for the environment. One way that ecologists test this is to place random collections of species in mini-environments and then measure some outcome. I have been working with a group of fresh-water ecologists to improve this in two ways. The first is that our subsets of species are carefully chosen, not random. The second is that we fit a nested family of plausible models. Our results suggest that the underlying model is not diversity at all.

Biodiversity experiments

This seems to be the received wisdom.

Treatments:	random sets of species
Measured response Y:	some eco-desirable outcome
Conclusion:	the greater the number of different species, the better the outcome.

A more carefully controlled experiment

A, B, C, D, E, F — six types of freshwater "shrimp".
Put 12 shrimps in a jar containing stream water and alder leaf litter.
Measure how much leaf litter is eaten after 28 days.

A more carefully controlled experiment

A, B, C, D, E, F — six types of freshwater “shrimp”.
Put 12 shrimps in a jar containing stream water and alder leaf litter.
Measure how much leaf litter is eaten after 28 days.

Experimental unit = jar.

A more carefully controlled experiment

A, B, C, D, E, F — six types of freshwater “shrimp”.
Put 12 shrimps in a jar containing stream water and alder leaf litter.
Measure how much leaf litter is eaten after 28 days.

Experimental unit = jar.

Treatment			Richness Level
A, . . . , F	monoculture	12 of type A	1
AB, . . . , EF	duoculture	6 of A, 6 of B	2
ABC, . . . , DEF	triculture	4 of A, 4 of B, 4 of C	3

A more carefully controlled experiment

A, B, C, D, E, F — six types of freshwater “shrimp”.
Put 12 shrimps in a jar containing stream water and alder leaf litter.
Measure how much leaf litter is eaten after 28 days.

Experimental unit = jar.

	Treatment			Richness Level
6	A, . . . , F	monoculture	12 of type A	1
15	AB, . . . , EF	duoculture	6 of A, 6 of B	2
20	ABC, . . . , DEF	triculture	4 of A, 4 of B, 4 of C	3
41				

A more carefully controlled experiment

A, B, C, D, E, F — six types of freshwater "shrimp".
Put 12 shrimps in a jar containing stream water and alder leaf litter.
Measure how much leaf litter is eaten after 28 days.

Experimental unit = jar.

	Treatment			Richness Level
6	A, . . . , F	monoculture	12 of type A	1
15	AB, . . . , EF	duoculture	6 of A, 6 of B	2
20	ABC, . . . , DEF	triculture	4 of A, 4 of B, 4 of C	3
41				

The experiment was carried out in 4 blocks of 41 jars.

Call in a statistician

The biologist fitted the model ‘Richness’ with 3 parameters,
one for each level of richness,
and found no evidence of any differences between the levels.

Call in a statistician

The biologist fitted the model ‘Richness’ with 3 parameters, one for each level of richness, and found no evidence of any differences between the levels.

I suggested the model ‘Type’ with 6 parameters $\alpha_A, \ldots, \alpha_F$:

monoculture A	α_A
duoculture AB	$\dfrac{\alpha_A + \alpha_B}{2}$
triculture ABC	$\dfrac{\alpha_A + \alpha_B + \alpha_C}{3}$

Call in a statistician

The biologist fitted the model 'Richness' with 3 parameters,
one for each level of richness,
and found no evidence of any differences between the levels.

I suggested the model 'Type' with 6 parameters $\alpha_A, \ldots, \alpha_F$:

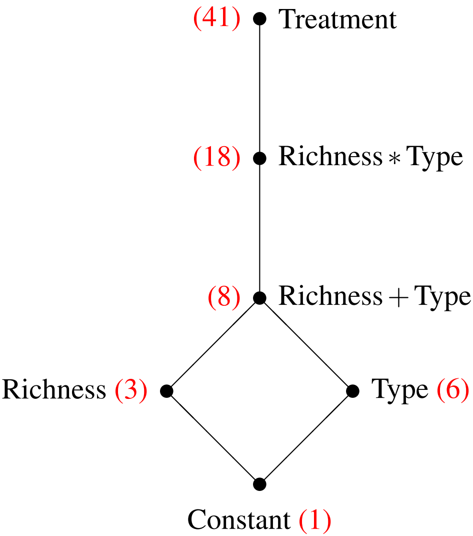
monoculture A	α_A
duoculture AB	$\dfrac{\alpha_A + \alpha_B}{2}$
triculture ABC	$\dfrac{\alpha_A + \alpha_B + \alpha_C}{3}$

In other words, if there are x_i shrimps of type i then

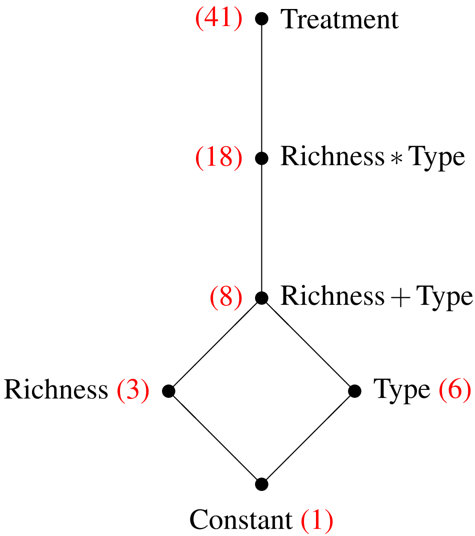
$$\mathbb{E}(Y) = \sum_{i=1}^{6} a_i x_i \qquad \text{where } 12a_i = \alpha_i.$$

($\sum x_i = 12$ always, so no need for intercept.)

Family of expectation models

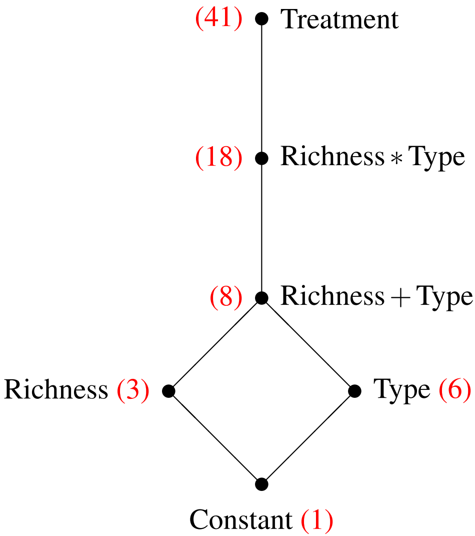


Family of expectation models



(add a different constant for each level of richness)

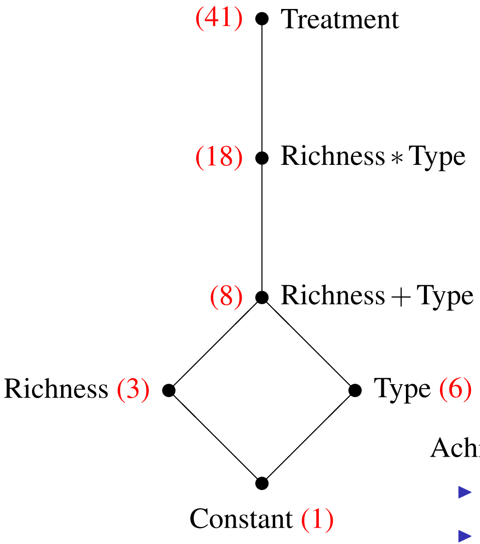
Family of expectation models



(a_i can change with each level of richness but does not depend on what else is present)

(add a different constant for each level of richness)

Family of expectation models



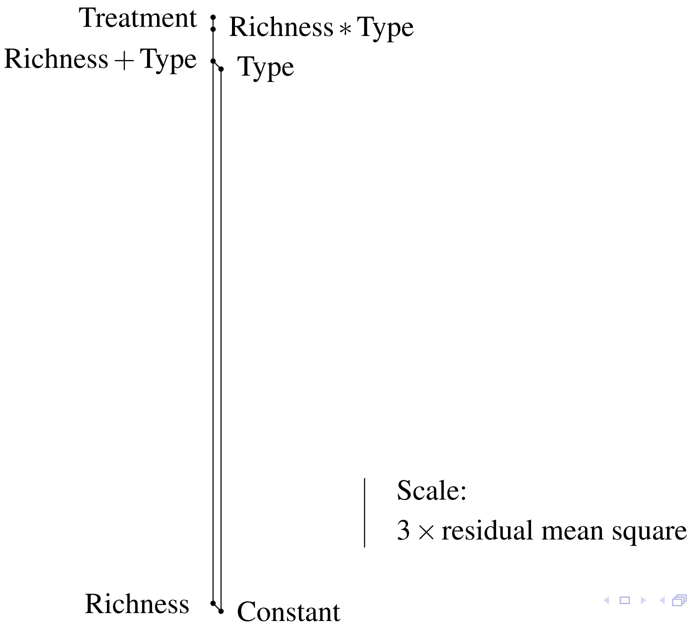
(a_i can change with each level of richness but does not depend on what else is present)

(add a different constant for each level of richness)

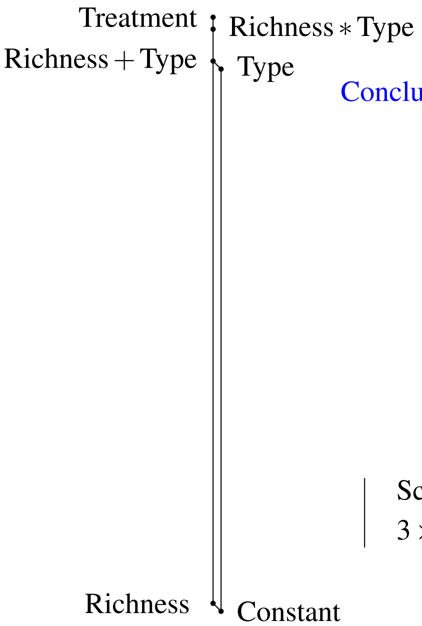
Achievement: an ecology journal published

- diagram of family of models
- statement that each row of an ANOVA table is for a difference between models.

What the data showed: mean squares



What the data showed: mean squares

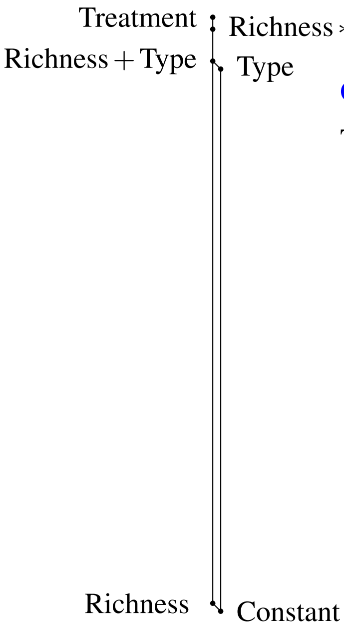


Conclusions:

Scale:

$3 \times$ residual mean square

What the data showed: mean squares



Conclusions:

The model Richness does not explain the data.

Scale:

3 × residual mean square

What the data showed: mean squares

Treatment

Richness $*$ Type

Richness $+$ Type

Type

Conclusions:

The model Richness does not explain the data.

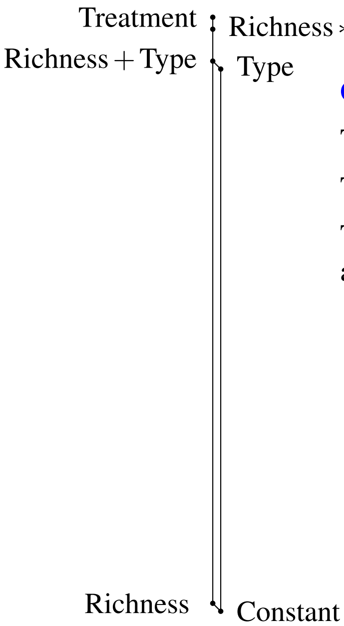
The model Type explains the data well.

Scale:

$3 \times$ residual mean square

Richness

Constant



Conclusions:

The model Richness does not explain the data.

The model Type explains the data well.

There is no evidence that any larger model does any better.

Scale:

$3 \times$ residual mean square

What the data showed: mean squares

Treatment
Richness + Type
Richness

Richness $*$ Type
Type
Constant

Conclusions:

The model Richness does not explain the data.

The model Type explains the data well.

There is no evidence that any larger model does any better.

Two experiments, with two responses each, all led to similar conclusions.

Scale:
$3 \times$ residual mean square

A new experiment on a different ecosystem (7 types)

Treatment			Richness Level
A, ..., G	monoculture	12 of type A	1
AB, ..., FG	duoculture	6 of A, 6 of B	2
ABC, ..., EFG	triculture	4 of A, 4 of B, 4 of C	3

A new experiment on a different ecosystem (7 types)

	Treatment			Richness Level
7	A, ..., G	monoculture	12 of type A	1
21	AB, ..., FG	duoculture	6 of A, 6 of B	2
35	ABC, ..., EFG	triculture	4 of A, 4 of B, 4 of C	3
63				

A new experiment on a different ecosystem (7 types)

	Treatment			Richness Level
7	A, ..., G	monoculture	12 of type A	1
21	AB, ..., FG	duoculture	6 of A, 6 of B	2
35	ABC, ..., EFG	triculture	4 of A, 4 of B, 4 of C	3
63				

"Do I really need all 35 tricultures?"

A new experiment on a different ecosystem (7 types)

	Treatment			Richness Level
7	A, ..., G	monoculture	12 of type A	1
21	AB, ..., FG	duoculture	6 of A, 6 of B	2
35	ABC, ..., EFG	triculture	4 of A, 4 of B, 4 of C	3
63				

"Do I really need all 35 tricultures?"

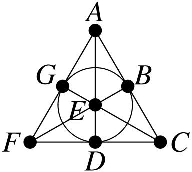
"Use 7 tricultures making a balanced incomplete-block design."

A new experiment on a different ecosystem (7 types)

	Treatment			Richness Level
7	A, …, G	monoculture	12 of type A	1
21	AB, …, FG	duoculture	6 of A, 6 of B	2
35	ABC, …, EFG	triculture	4 of A, 4 of B, 4 of C	3
63				

"Do I really need all 35 tricultures?"

"Use 7 tricultures making a balanced incomplete-block design."

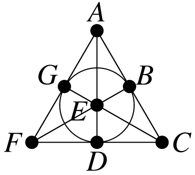


A new experiment on a different ecosystem (7 types)

	Treatment			Richness Level
7	A, ..., G	monoculture	12 of type A	1
21	AB, ..., FG	duoculture	6 of A, 6 of B	2
35	ABC, ..., EFG	triculture	4 of A, 4 of B, 4 of C	3
63				

"Do I really need all 35 tricultures?"

"Use 7 tricultures making a balanced incomplete-block design."



Another achievement:
an ecology journal published
a picture of the Fano plane.

How should such experiments be designed?

What is the purpose of the experiment?

How should such experiments be designed?

What is the purpose of the experiment?

- To find out whether the response changes with different levels of richness?

How should such experiments be designed?

What is the purpose of the experiment?

- To find out whether the response changes with different levels of richness?
- To estimate the differences between the different levels of richness?

How should such experiments be designed?

What is the purpose of the experiment?

- To find out whether the response changes with different levels of richness?
- To estimate the differences between the different levels of richness?
  - If so, we should replicate them more equally than $6 : 15 : 20$ or $7 : 21 : 7$.

How should such experiments be designed?

What is the purpose of the experiment?

- To find out whether the response changes with different levels of richness?
- To estimate the differences between the different levels of richness?
  - If so, we should replicate them more equally than $6 : 15 : 20$ or $7 : 21 : 7$.
- To discriminate between the (incomparable) models Richness and Type?

How should such experiments be designed?

What is the purpose of the experiment?

- To find out whether the response changes with different levels of richness?
- To estimate the differences between the different levels of richness?
  - If so, we should replicate them more equally than 6 : 15 : 20 or 7 : 21 : 7.
- To discriminate between the (incomparable) models Richness and Type?
- To discriminate between the model Type and the more general model which allows the response of each type to depend on what other types are present?

How should such experiments be designed?

What is the purpose of the experiment?

- To find out whether the response changes with different levels of richness?
- To estimate the differences between the different levels of richness?
  - If so, we should replicate them more equally than 6 : 15 : 20 or 7 : 21 : 7.
- To discriminate between the (incomparable) models Richness and Type?
- To discriminate between the model Type and the more general model which allows the response of each type to depend on what other types are present?
- To estimate the parameters (response per individual for each type) for the model Type?

How should such experiments be designed?

What is the purpose of the experiment?

- To find out whether the response changes with different levels of richness?
- To estimate the differences between the different levels of richness?
  - If so, we should replicate them more equally than $6:15:20$ or $7:21:7$.
- To discriminate between the (incomparable) models Richness and Type?
- To discriminate between the model Type and the more general model which allows the response of each type to depend on what other types are present?
- To estimate the parameters (response per individual for each type) for the model Type?
  - If so, we should not include any polycultures.

How should we choose which subsets to include?

Suppose that there are t types in all.
For a given level k of richness,
each treatment consists of equal numbers of each type in some subset of k types.
Suppose that we can include b such treatments.

How should we choose which subsets to include?

Suppose that there are t types in all.
For a given level k of richness,
each treatment consists of equal numbers of each type in some subset of k types.
Suppose that we can include b such treatments.

Should we

- choose subsets at random (traditional for ecologists)?

How should we choose which subsets to include?

Suppose that there are t types in all.
For a given level k of richness,
each treatment consists of equal numbers of each type in some subset of k types.
Suppose that we can include b such treatments.

Should we

- choose subsets at random (traditional for ecologists)?
- choose the subsets that would be best for an incomplete-block design with the usual linear model (e.g. the Fano plane)?

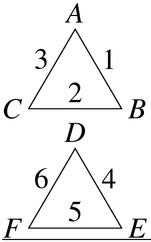
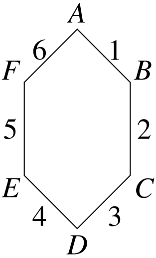
How should we choose which subsets to include?

Suppose that there are t types in all.
For a given level k of richness,
each treatment consists of equal numbers of each type in some subset of k types.
Suppose that we can include b such treatments.

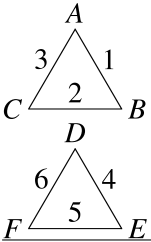
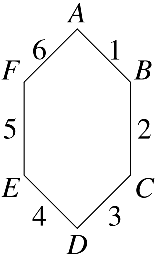
Should we

- choose subsets at random (traditional for ecologists)?
- choose the subsets that would be best for an incomplete-block design with the usual linear model (e.g. the Fano plane)?
- adopt some other strategy?

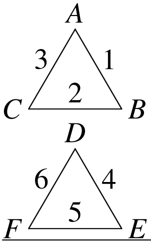
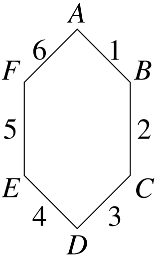
6 subsets of size 2, from 6 types

design		
		
		

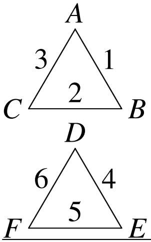
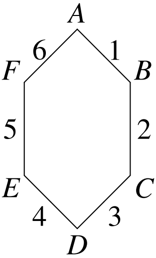
6 subsets of size 2, from 6 types

design	usual model 12 responses	
		
		

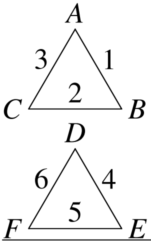
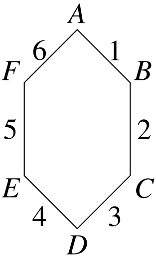
6 subsets of size 2, from 6 types

design	usual model 12 responses	
	$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$	
		

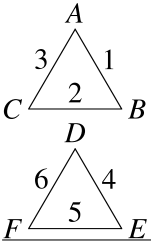
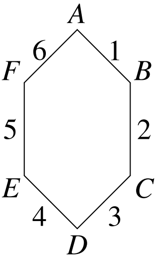
6 subsets of size 2, from 6 types

design	usual model 12 responses	
 	$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \dfrac{4\sigma^2}{3}$	
 		

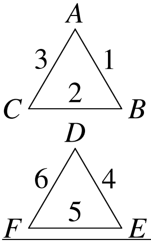
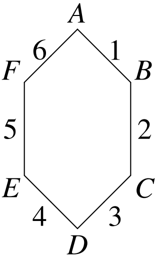
6 subsets of size 2, from 6 types

design	usual model 12 responses	
	$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \dfrac{4\sigma^2}{3}$ $\alpha_A - \alpha_D$ is not estimable	
		

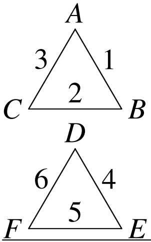
6 subsets of size 2, from 6 types

design	usual model 12 responses	duoculture 6 responses
	$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{4\sigma^2}{3}$ $\alpha_A - \alpha_D$ is not estimable	
		

6 subsets of size 2, from 6 types

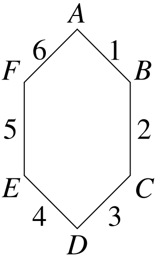
design	usual model 12 responses	duoculture 6 responses
	$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \dfrac{4\sigma^2}{3}$ $\alpha_A - \alpha_D$ is not estimable	$\mathbb{E}(Y_1) = \dfrac{\alpha_A + \alpha_B}{2}$
		

6 subsets of size 2, from 6 types

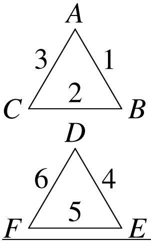
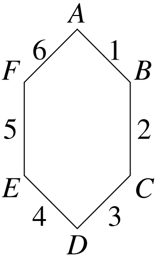
<table>
<tr><th>design</th><th>usual model
12 responses</th><th>duoculture
6 responses</th></tr>
<tr><td>
</td><td>$$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$$

$$\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{4\sigma^2}{3}$$

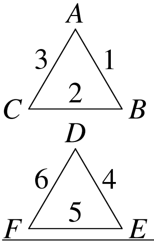
$\alpha_A - \alpha_D$ is not estimable</td><td>$$\mathbb{E}(Y_1) = \frac{\alpha_A + \alpha_B}{2}$$

$$\hat{\alpha}_A = Y_1 + Y_3 - Y_2$$</td></tr>
<tr><td>
</td><td></td><td></td></tr>
</table>

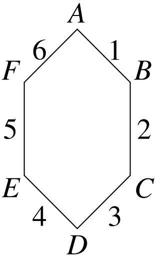
6 subsets of size 2, from 6 types

design	usual model 12 responses	duoculture 6 responses
	$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{4\sigma^2}{3}$ $\alpha_A - \alpha_D$ is not estimable	$\mathbb{E}(Y_1) = \frac{\alpha_A + \alpha_B}{2}$ $\hat{\alpha}_A = Y_1 + Y_3 - Y_2$ $\mathrm{Var}(\hat{\alpha}_A) = 3\sigma^2$
		

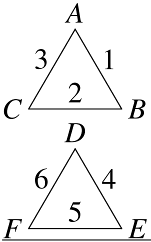
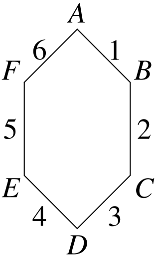
6 subsets of size 2, from 6 types

<table>
<tr><th>design</th><th>usual model
12 responses</th><th>duoculture
6 responses</th></tr>
<tr><td>
</td><td>$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$

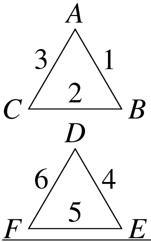
$\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{4\sigma^2}{3}$

$\alpha_A - \alpha_D$ is not estimable</td><td>$\mathbb{E}(Y_1) = \frac{\alpha_A + \alpha_B}{2}$
$\hat{\alpha}_A = Y_1 + Y_3 - Y_2$
$\mathrm{Var}(\hat{\alpha}_A) = 3\sigma^2$
$\sum_i \mathrm{Var}(\hat{\alpha}_i)$ is minimized</td></tr>
<tr><td>
</td><td></td><td></td></tr>
</table>

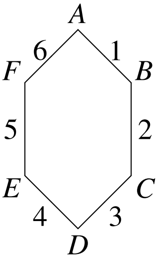
6 subsets of size 2, from 6 types

design	usual model 12 responses	duoculture 6 responses
	$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{4\sigma^2}{3}$ $\alpha_A - \alpha_D$ is not estimable	$\mathbb{E}(Y_1) = \frac{\alpha_A + \alpha_B}{2}$ $\hat{\alpha}_A = Y_1 + Y_3 - Y_2$ $\mathrm{Var}(\hat{\alpha}_A) = 3\sigma^2$ $\sum_i \mathrm{Var}(\hat{\alpha}_i)$ is minimized
	$\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_D) = 3\sigma^2$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_C) = \frac{8\sigma^2}{3}$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{5\sigma^2}{3}$	

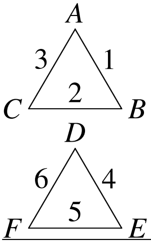
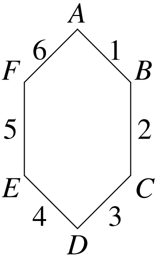
6 subsets of size 2, from 6 types

<table>
<tr><th>design</th><th>usual model
12 responses</th><th>duoculture
6 responses</th></tr>
<tr><td></td><td>$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$

$\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{4\sigma^2}{3}$

$\alpha_A - \alpha_D$ is not estimable</td><td>$\mathbb{E}(Y_1) = \frac{\alpha_A + \alpha_B}{2}$
$\hat{\alpha}_A = Y_1 + Y_3 - Y_2$
$\mathrm{Var}(\hat{\alpha}_A) = 3\sigma^2$
$\sum_i \mathrm{Var}(\hat{\alpha}_i)$ is minimized</td></tr>
<tr><td></td><td>$\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_D) = 3\sigma^2$
$\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_C) = \frac{8\sigma^2}{3}$
$\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{5\sigma^2}{3}$
$\sum_{i \neq j} \mathrm{Var}(\hat{\alpha}_i - \hat{\alpha}_j)$ is minimized</td><td></td></tr>
</table>

6 subsets of size 2, from 6 types

design	usual model 12 responses	duoculture 6 responses
	$\mathbb{E}(Y_{1A}) = \alpha_A + \beta_B$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{4\sigma^2}{3}$ $\alpha_A - \alpha_D$ is not estimable	$\mathbb{E}(Y_1) = \frac{\alpha_A + \alpha_B}{2}$ $\hat{\alpha}_A = Y_1 + Y_3 - Y_2$ $\mathrm{Var}(\hat{\alpha}_A) = 3\sigma^2$ $\sum_i \mathrm{Var}(\hat{\alpha}_i)$ is minimized
	$\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_D) = 3\sigma^2$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_C) = \frac{8\sigma^2}{3}$ $\mathrm{Var}(\hat{\alpha}_A - \hat{\alpha}_B) = \frac{5\sigma^2}{3}$ $\sum_{i \neq j} \mathrm{Var}(\hat{\alpha}_i - \hat{\alpha}_j)$ is minimized	α_A is not estimable

A non-intuitive result

Consider incomplete-block designs for t treatments in b blocks of size k.

Usual model	expected response on any unit with treatment i in block B is $\alpha_i + \beta_B$
Polyculture model	expected response on any unit whose treatment is an equal mixture of the species in subset B is $\sum_{i \in B} \alpha_i / k$

A non-intuitive result

Consider incomplete-block designs for t treatments in b blocks of size k.

Usual model	expected response on any unit with treatment i in block B is $\alpha_i + \beta_B$
Polyculture model	expected response on any unit whose treatment is an equal mixture of the species in subset B is $\sum_{i \in B} \alpha_i / k$

If there is no balanced incomplete-block design for t treatments in b blocks of size k then a design which is best for one situation may be worst for the other.

References

1. Julia Reiss, R. A. Bailey, Fernanda Cássio, Guy Woodward and Cláudia Pascoal: Assessing the contribution of micro-organisms and macrofauna to biodiversity-ecosystem functioning relationships in freshwater microcosms. *Advances in Ecological Research*, **43**, 2010, 151–176.
2. Julia Reiss, R. A. Bailey, Daniel M. Perkins, Angela Pluchinotta and Guy Woodward: Testing effects of consumer richness, evenness and body size on ecosystem functioning. *Journal of Animal Ecology*, 2011, in press.