Design and analysis of experiments testing for biodiversity

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

Theoretical advances in experimental design
Royal Statistical Society 2019 International Conference
Belfast, September 2019

Ongoing joint work with Julia Reiss and Daniel Perkins

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. Others use a carefully chosen collection of species.

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. Others use a carefully chosen collection of species.

Is the outcome affected by the number of different species present, or is it just the number of members of each species that matters? And is there an interaction between these? Are the answers affected by the temperature, or by the complexity of the environment?

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. Others use a carefully chosen collection of species.

Is the outcome affected by the number of different species present,
or is it just the number of members of each species that matters?
And is there an interaction between these? Are the answers affected by the temperature, or by the complexity of the environment?

I have been working with a group of fresh-water ecologists on the design and analysis of such experiments. Our subsets of species are carefully chosen, not random. We design the combinations of these subsets with levels of other factors.

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. Others use a carefully chosen collection of species.

Is the outcome affected by the number of different species present,
or is it just the number of members of each species that matters?
And is there an interaction between these? Are the answers affected by the temperature, or by the complexity of the environment?

I have been working with a group of fresh-water ecologists on the design and analysis of such experiments. Our subsets of species are carefully chosen, not random. We design the combinations of these subsets with levels of other factors.

We also fit a nested family of plausible models to the data.
Our results suggest that the underlying model is not diversity at all. One of my crucial inputs has been the use of Hasse diagrams as a way of understanding a complicated family of plausible models for the expectation of the response.

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.

Assemblage identity			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.

	Assemblage identity			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.

	Assemblage identity			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.

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.

	Assemblage identity			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.
Actually 42 jars, because untreated jars were included,
but their data was so obviously different that it was excluded from further modelling.

Initial model fitting

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.

Initial model fitting

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.

This model for the response Y is

$$\mathbb{E}(Y) = \begin{cases} \alpha_1 & \text{on monocultures A, \ldots, F} \\ \alpha_2 & \text{on duocultures AB, \ldots, EF} \\ \alpha_3 & \text{on tricultures ABC, \ldots, DEF} \end{cases}$$

Initial model fitting

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.

This model for the response Y is

$$\mathbb{E}(Y) = \begin{cases} \alpha_1 & \text{on monocultures A}, \ldots, \text{F} \\ \alpha_2 & \text{on duocultures AB}, \ldots, \text{EF} \\ \alpha_3 & \text{on tricultures ABC}, \ldots, \text{DEF} \end{cases}$$

The data did not give any evidence against the null hypothesis that

$$\alpha_1 = \alpha_2 = \alpha_3 :$$

this is the ‘Constant’ model, or null model.

Call in a statistician

		Assemblage identity	R	$x1$	$x2$	$x3$	$x4$	$x5$	$x6$
1	A	12 of type A	1	12	0	0	0	0	0
⋮			⋮						
6	F	12 of type F	1	0	0	0	0	0	12
7	AB	6 of A, 6 of B	2	6	6	0	0	0	0
⋮			⋮						
21	EF	6 of E, 6 of F	2	0	0	0	0	6	6
22	ABC	4 of A, 4 of B, 4 of C	3	4	4	4	0	0	0
⋮			⋮						
41	DEF	4 of D, 4 of E, 4 of F	3	0	0	0	4	4	4

Call in a statistician

Assemblage identity			R	$x1$	$x2$	$x3$	$x4$	$x5$	$x6$
1	A	12 of type A	1	12	0	0	0	0	0
⋮			⋮						
6	F	12 of type F	1	0	0	0	0	0	12
7	AB	6 of A, 6 of B	2	6	6	0	0	0	0
⋮			⋮						
21	EF	6 of E, 6 of F	2	0	0	0	0	6	6
22	ABC	4 of A, 4 of B, 4 of C	3	4	4	4	0	0	0
⋮			⋮						
41	DEF	4 of D, 4 of E, 4 of F	3	0	0	0	4	4	4

I suggested the model 'Type' with 6 parameters $\beta_1, \ldots, \beta_6$:

$$\mathbb{E}(Y) = \sum_{i=1}^{6} \beta_i x_i$$

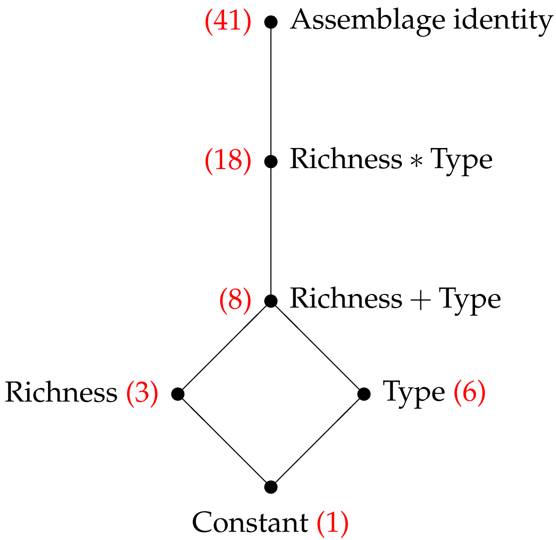
Call in a statistician

		Assemblage identity	R	$x1$	$x2$	$x3$	$x4$	$x5$	$x6$
1	A	12 of type A	1	12	0	0	0	0	0
$\vdots$			$\vdots$						
6	F	12 of type F	1	0	0	0	0	0	12
7	AB	6 of A, 6 of B	2	6	6	0	0	0	0
$\vdots$			$\vdots$						
21	EF	6 of E, 6 of F	2	0	0	0	0	6	6
22	ABC	4 of A, 4 of B, 4 of C	3	4	4	4	0	0	0
$\vdots$			$\vdots$						
41	DEF	4 of D, 4 of E, 4 of F	3	0	0	0	4	4	4

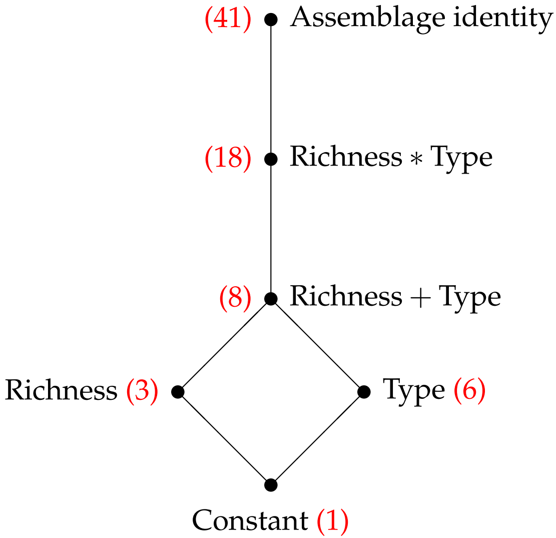
I suggested the model ‘Type’ with 6 parameters $\beta_1, \ldots, \beta_6$:

$$\mathbb{E}(Y) = \sum_{i=1}^{6} \beta_i x_i \qquad \left(\sum x_i = 12 \text{ always, so no need for intercept.}\right)$$

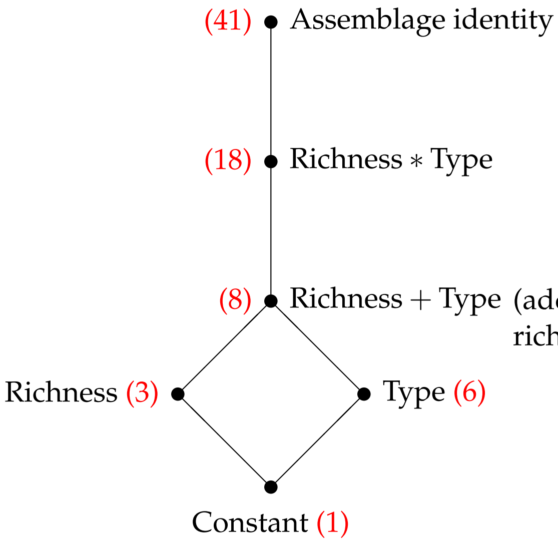
Family of expectation models (subspaces): dimensions shown in red



Family of expectation models (subspaces): dimensions shown in red

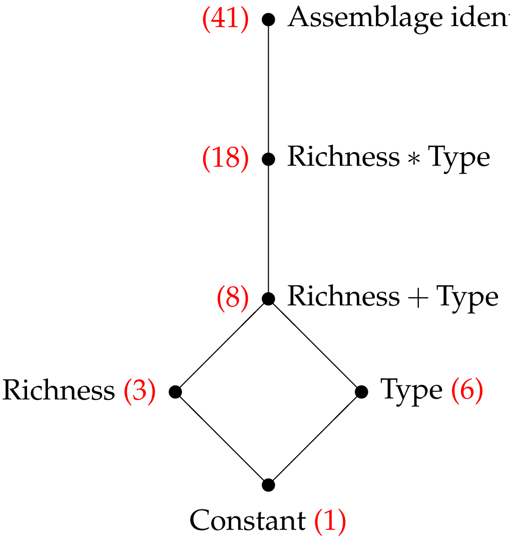


The subspace at the bottom of a line is contained in the subspace at the top of it



The subspace at the bottom of a line is contained in the subspace at the top of it

Richness + Type (add a different constant for each level of richness)

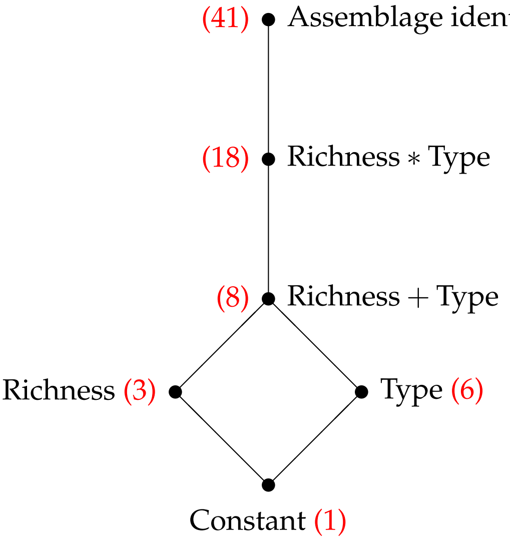


The subspace at the bottom of a line is contained in the subspace at the top of it

Richness ∗ Type: (β_i can change with each level of richness but does not depend on what else is present)

Richness + Type: (add a different constant for each level of richness)

Family of expectation models (subspaces): dimensions shown in red



The subspace at the bottom of a line is contained in the subspace at the top of it

Richness ∗ Type: (β_i can change with each level of richness but does not depend on what else is present)

Richness + Type: (add a different constant for each level of richness)

Success: an ecology journal published

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

Analysis of Variance (ANOVA) table

Source	df	SS	MS	F	P
Richness	2	0.000009	0.000005	0.49	n.s.
Type	5	0.003859	0.000772	81.37	< 0.0005
Richness $*$ Type	10	0.000127	0.000013	1.34	n.s.
Assemblage Identity	23	0.000105	0.000005	0.48	n.s.
Block	3	0.000067	0.000022		
Error	120	0.001138	0.000009		
Total	163	0.005306			

Analysis of Variance (ANOVA) table

Source	df	SS	MS	F	P
Richness	2	0.000009	0.000005	0.49	n.s.
Type	5	0.003859	0.000772	81.37	< 0.0005
Richness $*$ Type	10	0.000127	0.000013	1.34	n.s.
Assemblage Identity	23	0.000105	0.000005	0.48	n.s.
Block	3	0.000067	0.000022		
Error	120	0.001138	0.000009		
Total	163	0.005306			

Each row in the ANOVA table represents not a model but the difference between a larger model and the next smaller one. See Fig. 1 for how the models are related.

Analysis of Variance (ANOVA) table

Source	df	SS	MS	F	P
Richness	2	0.000009	0.000005	0.49	n.s.
Type	5	0.003859	0.000772	81.37	< 0.0005
Richness $*$ Type	10	0.000127	0.000013	1.34	n.s.
Assemblage Identity	23	0.000105	0.000005	0.48	n.s.
Block	3	0.000067	0.000022		
Error	120	0.001138	0.000009		
Total	163	0.005306			

Each row in the ANOVA table represents not a model but the difference between a larger model and the next smaller one. See Fig. 1 for how the models are related.

Verbatim from *Journal of Animal Ecology*

Scaling the Hasse diagram of expectation subspaces

Suppose that V_1 and V_2 are expectation subspaces,
with $V_1 < V_2$,
and an edge joining V_1 to V_2.

Scaling the Hasse diagram of expectation subspaces

Suppose that V_1 and V_2 are expectation subspaces,
with $V_1 < V_2$,
and an edge joining V_1 to V_2.

The mean square for
the extra fit in V_2 compared to the fit in V_1 is

$$\frac{\text{SS(fitted values in } V_2) \; - \; \text{SS(fitted values in } V_1)}{\dim(V_2) - \dim(V_1)}.$$

Scaling the Hasse diagram of expectation subspaces

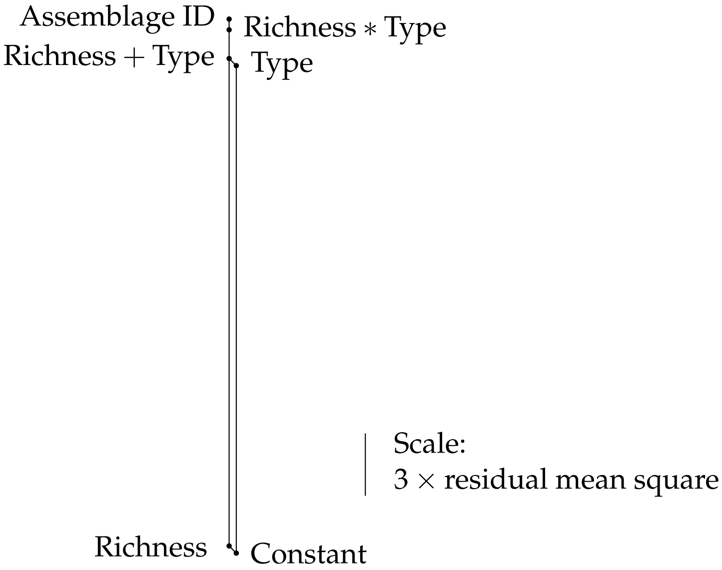
Suppose that V_1 and V_2 are expectation subspaces,
with $V_1 < V_2$,
and an edge joining V_1 to V_2.

The mean square for
the extra fit in V_2 compared to the fit in V_1 is

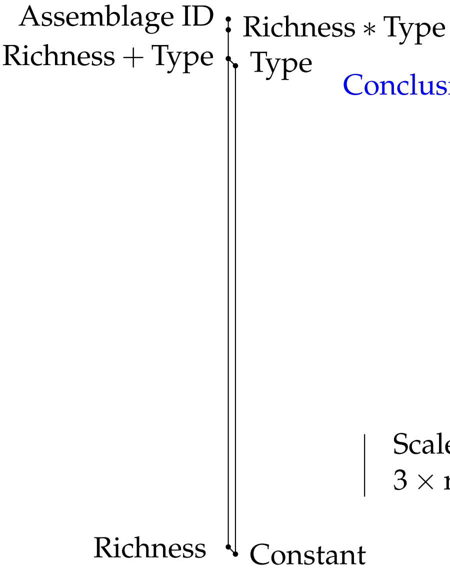
$$\frac{\text{SS(fitted values in } V_2) \; - \; \text{SS(fitted values in } V_1)}{\dim(V_2) - \dim(V_1)}.$$

Scale the Hasse diagram so that each edge has length proportional to the relevant mean square,
and show the residual mean square to give a scale.

What the data showed: mean squares

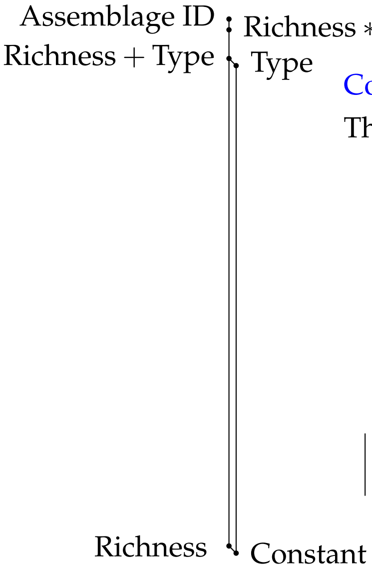


What the data showed: mean squares



Conclusions:

Scale:
3 × residual mean square

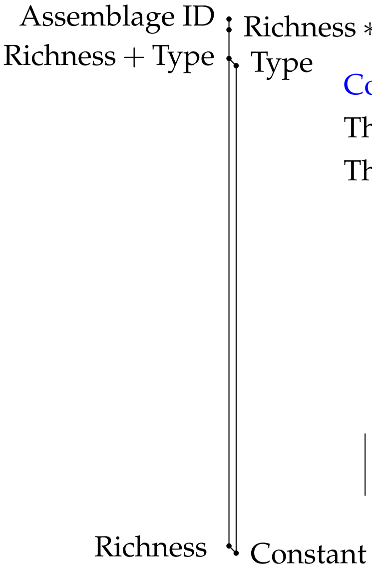


Conclusions:

The model Richness does not explain the data.

Scale:
3 × residual mean square

What the data showed: mean squares



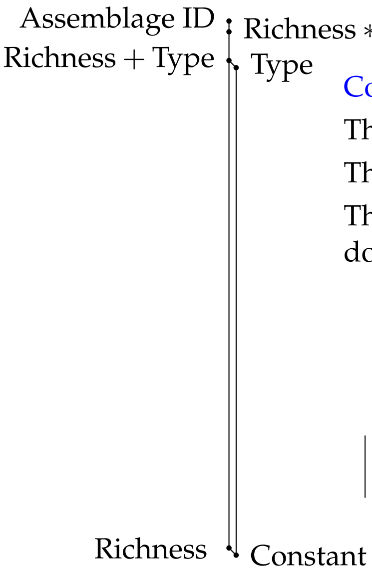
Conclusions:

The model Richness does not explain the data.

The model Type explains the data well.

Scale:
3 × residual mean square

What the data showed: mean squares



Conclusions:

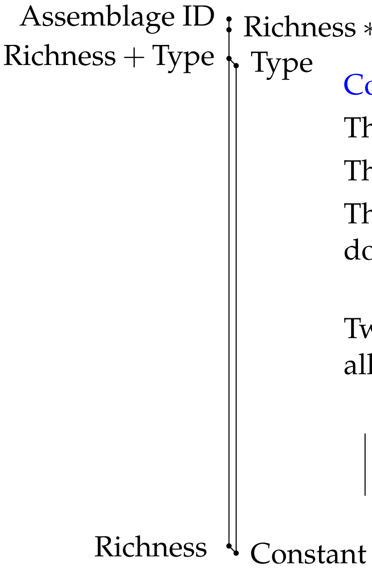
The model Richness does not explain the data.

The model Type explains the data well.

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

Scale:
3 × residual mean square

What the data showed: mean squares



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

Communicating the ideas

For complicated families of models,
the biologists that I work with often find the Hasse diagram
easier to understand than the equations.

Communicating the ideas

For complicated families of models,
the biologists that I work with often find the Hasse diagram
easier to understand than the equations.

Similarly, they find the scaled Hasse diagram
easier to understand than the anova table.

A new experiment on a different ecosystem (7 types)

Assemblage identity			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)

	Assemblage identity			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)

	Assemblage identity			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)

	Assemblage identity			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?"

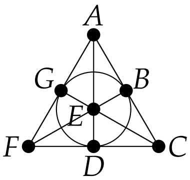
"Use 7 tricultures making a balanced incomplete-block design."

A new experiment on a different ecosystem (7 types)

	Assemblage identity			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."

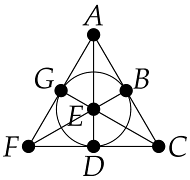


A new experiment on a different ecosystem (7 types)

	Assemblage identity			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 success: *Advances in Ecological Research* published this 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 n 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 n 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)?
  (b subsets, each used the same number of times)

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 n 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)?
  (b subsets, each used the same number of times)
- choose the subsets that would be best for an incomplete-block design with the usual linear model
  (e.g. the Fano plane)?
  (b blocks, each of k plots; one type on each plot;
  each plot measured once)

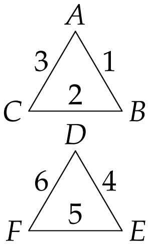
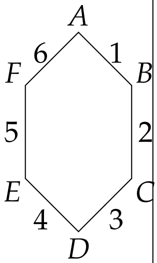
How should we choose which subsets to include?

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

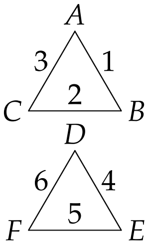
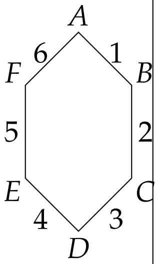
Should we

- choose subsets at random (traditional for ecologists)?
  (b subsets, each used the same number of times)
- choose the subsets that would be best for an incomplete-block design with the usual linear model
  (e.g. the Fano plane)?
  (b blocks, each of k plots; one type on each plot;
  each plot measured once)
- adopt some other strategy?

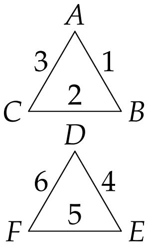
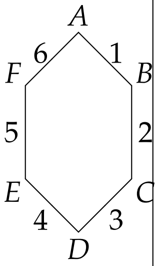
6 subsets of size 2, from 6 types

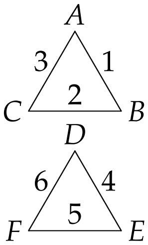
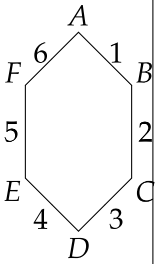
design		
		
		

6 subsets of size 2, from 6 types

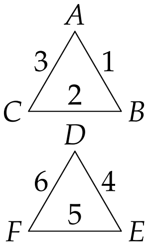
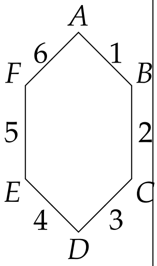
design	usual IBD model: 12 responses	
		
		

6 subsets of size 2, from 6 types

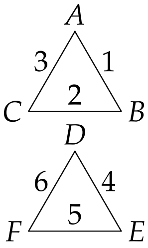
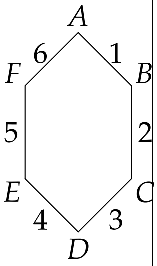
design	usual IBD model: 12 responses	
	$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$	
		

design	usual IBD model: 12 responses	
	$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$ $\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{4\sigma^2}{3}$	
		

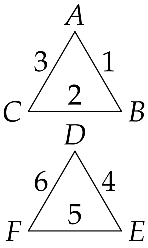
6 subsets of size 2, from 6 types

<table>
<tr><th>design</th><th>usual IBD model: 12 responses</th><th></th></tr>
<tr><td>
</td><td>$$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$$
$$\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{4\sigma^2}{3}$$
$\tau_A - \tau_D$ is not estimable</td><td></td></tr>
<tr><td>
</td><td></td><td></td></tr>
</table>

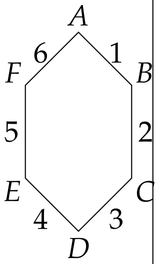
6 subsets of size 2, from 6 types

design	usual IBD model: 12 responses	duocultures: 6 responses
	$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$ $\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \dfrac{4\sigma^2}{3}$ $\tau_A - \tau_D$ is not estimable	
		

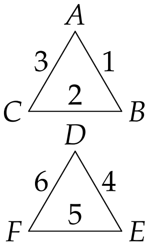
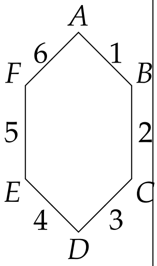
6 subsets of size 2, from 6 types

<table>
<tr><th>design</th><th>usual IBD model: 12 responses</th><th>duocultures: 6 responses</th></tr>
<tr><td>
</td><td>$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$

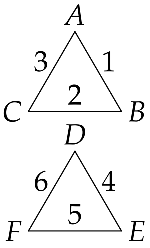
$\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{4\sigma^2}{3}$

$\tau_A - \tau_D$ is not estimable</td><td>$\mathbb{E}(Y_1) = n(\beta_A + \beta_B)$</td></tr>
<tr><td>
</td><td></td><td></td></tr>
</table>

6 subsets of size 2, from 6 types

<table>
<tr><th>design</th><th>usual IBD model: 12 responses</th><th>duocultures: 6 responses</th></tr>
<tr><td></td><td>$$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$$
$$\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{4\sigma^2}{3}$$
$\tau_A - \tau_D$ is not estimable</td><td>$$\mathbb{E}(Y_1) = n(\beta_A + \beta_B)$$
$$\hat{\beta}_A = \frac{Y_1 + Y_3 - Y_2}{2n}$$</td></tr>
<tr><td></td><td></td><td></td></tr>
</table>

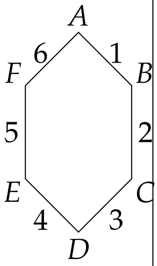
6 subsets of size 2, from 6 types

<table>
<tr><th>design</th><th>usual IBD model: 12 responses</th><th>duocultures: 6 responses</th></tr>
<tr><td>
</td><td>$$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$$

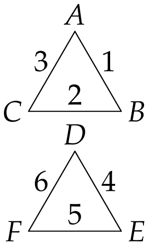
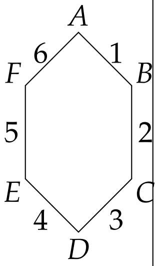
$$\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{4\sigma^2}{3}$$

$\tau_A - \tau_D$ is not estimable</td><td>$$\mathbb{E}(Y_1) = n(\beta_A + \beta_B)$$

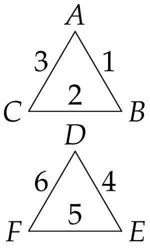
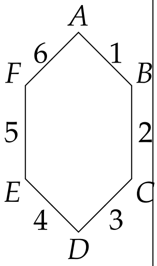
$$\hat{\beta}_A = \frac{Y_1 + Y_3 - Y_2}{2n}$$

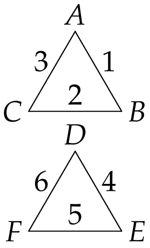
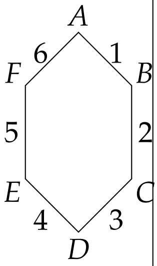
$$\mathrm{Var}(\hat{\beta}_A) = \frac{3\sigma^2}{4n^2}$$</td></tr>
<tr><td>
</td><td></td><td></td></tr>
</table>

6 subsets of size 2, from 6 types

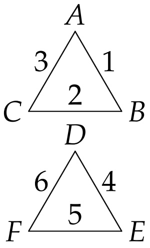
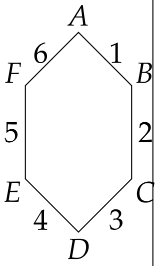
design	usual IBD model: 12 responses	duocultures: 6 responses
	$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$ $\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{4\sigma^2}{3}$ $\tau_A - \tau_D$ is not estimable	$\mathbb{E}(Y_1) = n(\beta_A + \beta_B)$ $\hat{\beta}_A = \frac{Y_1 + Y_3 - Y_2}{2n}$ $\mathrm{Var}(\hat{\beta}_A) = \frac{3\sigma^2}{4n^2}$ $\sum_i \mathrm{Var}(\hat{\beta}_i)$ is minimized
		

6 subsets of size 2, from 6 types

design	usual IBD model: 12 responses	duocultures: 6 responses
	$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$ $\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{4\sigma^2}{3}$ $\tau_A - \tau_D$ is not estimable	$\mathbb{E}(Y_1) = n(\beta_A + \beta_B)$ $\hat{\beta}_A = \frac{Y_1 + Y_3 - Y_2}{2n}$ $\mathrm{Var}(\hat{\beta}_A) = \frac{3\sigma^2}{4n^2}$ $\sum_i \mathrm{Var}(\hat{\beta}_i)$ is minimized
	$\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_D) = 3\sigma^2$ $\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_C) = \frac{8\sigma^2}{3}$ $\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{5\sigma^2}{3}$	

<table>
<tr><th>design</th><th>usual IBD model: 12 responses</th><th>duocultures: 6 responses</th></tr>
<tr>
<td></td>
<td>$$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$$ $$\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{4\sigma^2}{3}$$ $\tau_A - \tau_D$ is not estimable</td>
<td>$$\mathbb{E}(Y_1) = n(\beta_A + \beta_B)$$ $$\hat{\beta}_A = \frac{Y_1 + Y_3 - Y_2}{2n}$$ $$\mathrm{Var}(\hat{\beta}_A) = \frac{3\sigma^2}{4n^2}$$ $\sum_i \mathrm{Var}(\hat{\beta}_i)$ is minimized</td>
</tr>
<tr>
<td></td>
<td>$$\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_D) = 3\sigma^2$$ $$\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_C) = \frac{8\sigma^2}{3}$$ $$\mathrm{Var}(\hat{\tau}_A - \hat{\tau}_B) = \frac{5\sigma^2}{3}$$ $\sum_{i \neq j} \mathrm{Var}(\hat{\tau}_i - \hat{\tau}_j)$ is minimized</td>
<td></td>
</tr>
</table>

6 subsets of size 2, from 6 types

design	usual IBD model: 12 responses	duocultures: 6 responses
	$\mathbb{E}(Y_{1A}) = \tau_A + \delta_1$ $\mathrm{Var}(\hat\tau_A - \hat\tau_B) = \frac{4\sigma^2}{3}$ $\tau_A - \tau_D$ is not estimable	$\mathbb{E}(Y_1) = n(\beta_A + \beta_B)$ $\hat\beta_A = \frac{Y_1 + Y_3 - Y_2}{2n}$ $\mathrm{Var}(\hat\beta_A) = \frac{3\sigma^2}{4n^2}$ $\sum_i \mathrm{Var}(\hat\beta_i)$ is minimized
	$\mathrm{Var}(\hat\tau_A - \hat\tau_D) = 3\sigma^2$ $\mathrm{Var}(\hat\tau_A - \hat\tau_C) = \frac{8\sigma^2}{3}$ $\mathrm{Var}(\hat\tau_A - \hat\tau_B) = \frac{5\sigma^2}{3}$ $\sum_{i \neq j} \mathrm{Var}(\hat\tau_i - \hat\tau_j)$ is minimized	β_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 $\tau_i + \delta_B$
Polyculture model	expected response on any unit with a collection of n individuals from each species in subset B is $n \sum_{i \in B} \beta_i$

A non-intuitive result

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

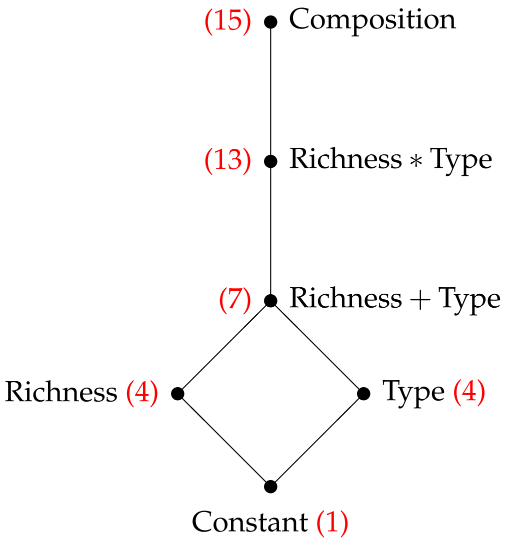
Usual model	expected response on any unit with treatment i in block B is $\tau_i + \delta_B$
Polyculture model	expected response on any unit with a collection of n individuals from each species in subset B is $n \sum_{i \in B} \beta_i$

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.

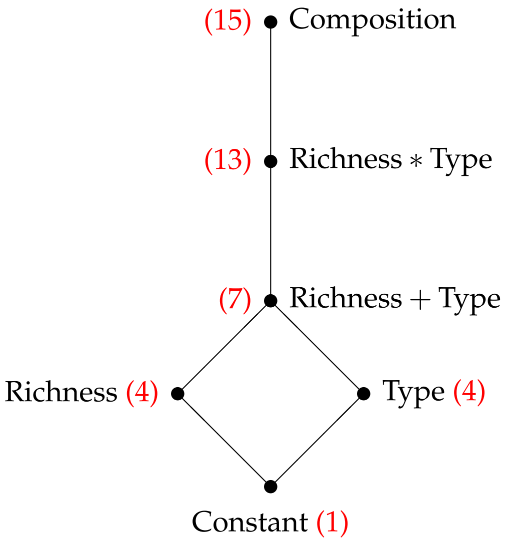
One aspect of a third biodiversity experiment: 4 types of "shrimp"

	Composition		Richness	$x1$	$x2$	$x3$	$x4$
1	A	12 of type A	1	12	0	0	0
2	B	12 of type B	1	0	12	0	0
3	C	12 of type C	1	0	0	12	0
4	D	12 of type D	1	0	0	0	12
5	AB	6 of A, 6 of B	2	6	6	0	0
6	AC	6 of A, 6 of C	2	6	0	6	0
7	AD	6 of A, 6 of D	2	6	0	0	6
8	BC	6 of B, 6 of C	2	0	6	6	0
9	BD	6 of B, 6 of D	2	0	6	0	6
10	CD	6 of C, 6 of D	2	0	0	6	6
11	ABC	4 of A, 4 of B, 4 of C	3	4	4	4	0
12	ABD	4 of A, 4 of B, 4 of D	3	4	4	0	4
13	ACD	4 of A, 4 of C, 4 of D	3	4	0	4	4
14	BCD	4 of B, 4 of C, 4 of D	3	0	4	4	4
15	$ABCD$	3 each of A, B, C and D	4	3	3	3	3

Family of expectation models (so far)



Family of expectation models (so far)



Five responses were measured. For every response, the sum of squares of fitted values for Composition was hardly any bigger than the sum of squares of fitted values for the model Richness ∗ Type, so we decided to omit Richness ∗ Type.

Other details of the third experiment

Each of the 15 compositions was combined with three temperatures: 5^o C, 10^o C and 15^oC.

Other details of the third experiment

Each of the 15 compositions was combined with three temperatures: 5^o C, 10^o C and $15^o C$.

Each of the 45 combinations was replicated twice.

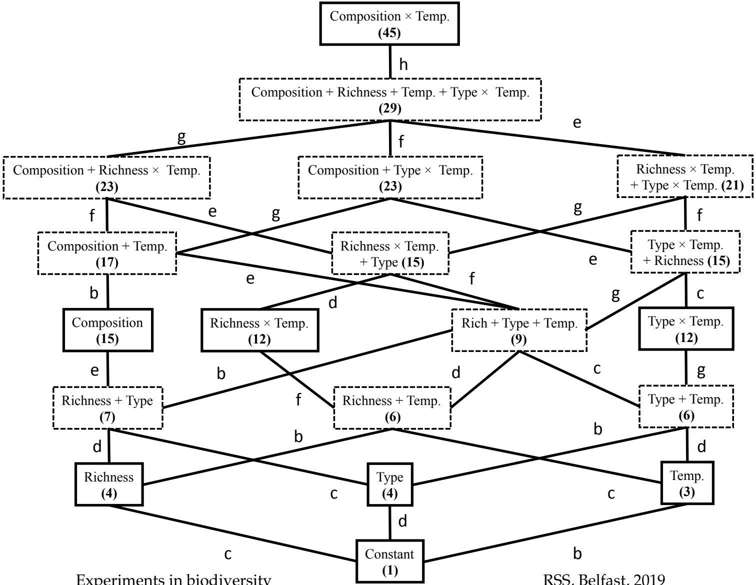
Other details of the third experiment

Each of the 15 compositions was combined with
three temperatures: 5^o C, 10^o C and 15^oC.

Each of the 45 combinations was replicated twice.

Three temperature-controlled rooms in a lab were used.
Each room had a single temperature and two of each composition.
Therefore there was no appropriate residual mean square to compare the main effect of Temperature with, but all other effects could be assessed.

Diagram from a paper in *Global Change Biology*



Brief results from the third biodiversity experiment

For each single type of response,
Type $\times$ Temperature explained the data well,
with no need for further terms.

Brief results from the third biodiversity experiment

For each single type of response,
Type $\times$ Temperature explained the data well,
with no need for further terms.

For multifunctionality, for each of the five types of response,
the mean of the three best outcomes was calculated.

Brief results from the third biodiversity experiment

For each single type of response,
Type $\times$ Temperature explained the data well,
with no need for further terms.

For multifunctionality, for each of the five types of response,
the mean of the three best outcomes was calculated.
For each of the 45 treatment combinations,
we recorded the number of types of response
on which the mean outcome exceeded 25% of this "best score".
On this measure, compositions with high levels of Richness scored well.

Brief results from the third biodiversity experiment

For each single type of response,
Type $\times$ Temperature explained the data well,
with no need for further terms.

For multifunctionality, for each of the five types of response,
the mean of the three best outcomes was calculated.
For each of the 45 treatment combinations,
we recorded the number of types of response
on which the mean outcome exceeded 25% of this “best score”.
On this measure, compositions with high levels of Richness scored well.

Note that this is a simple consequence of the model

$$\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

if the rankings of β_1, β_2, β_3 and β_4 are different over the five types of response.

One aspect of a fourth biodiversity experiment

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

Experimental unit = jar.

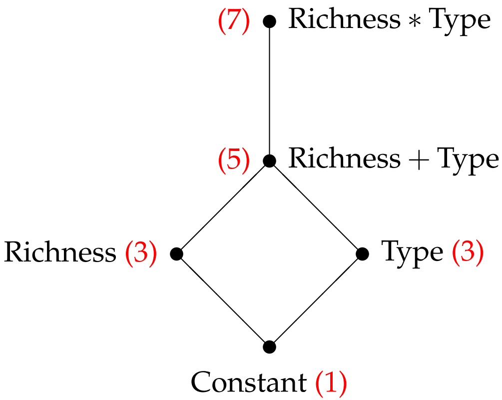
One aspect of a fourth biodiversity experiment

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

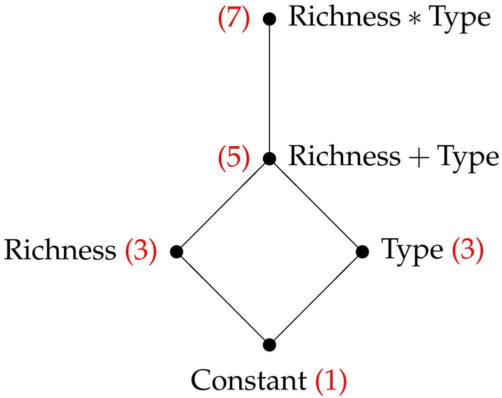
Experimental unit = jar.

	Assemblage identity		Richness	$x1$	$x2$	$x3$
1	A	12 of type A	1	12	0	0
2	B	12 of type B	1	0	12	0
3	C	12 of type C	1	0	0	12
4	AB	6 of A, 6 of B	2	6	6	0
5	AC	6 of A, 6 of C	2	6	0	6
6	BC	6 of B, 6 of C	2	0	6	6
7	ABC	4 of A, 4 of B, 4 of C	3	4	4	4

Family of expectation models (so far)



Family of expectation models (so far)



For these numbers, Assemblage identity = Richness ∗ Type.

The other aspect of the biodiversity experiment

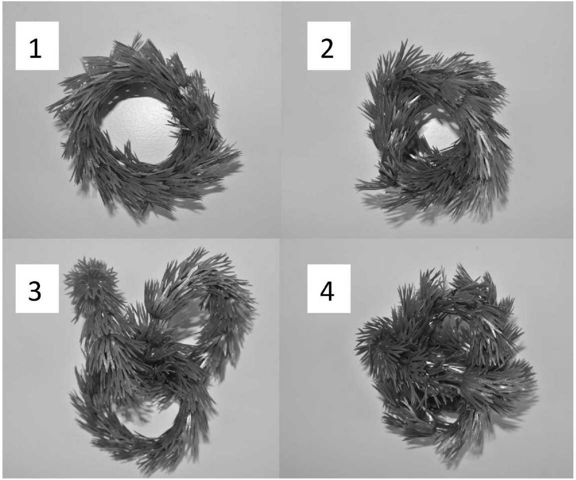


Fig 1. Photographs of the structures used to create habitat complexity in microcosms with 'structure present'. The basic unit of each structure was a plastic plant strip (mimicking *Ceratophyllum* spp.), joined up as a ring (~ 8cm in diameter) and four levels of fractal dimension were created with them: 1) level 1 consisted of two rings aligned, with a fractal dimension (D) of 1.77; 2) level 2 consisted of two rings twisted into each other (D = 1.80); 3) level 3 consisted of three rings locked together (D = 1.81) and 4) level four was a ball made from 3 rings together (D = 1.83). This design therefore also gave two levels of 'amount of structure' - 3 g for complexity level 1 and 2 and 4.5 g for complexity level 3 and 4.

Hasse diagram for enviromental model subspaces

(5) Fractal dimension (how twiddly were the extra bits?)

(3) Number of plastic rings (0 or 2 or 3)

(2) Plastic rings or not

(1) Constant

The experiment: 3 blocks, each with 35 jars

Environment Complexity	Assemblage identity						
	A	*B*	*C*	*AB*	*AC*	*BC*	*ABC*
0	×	×	×	×	×	×	×
1	×	×	×	×	×	×	×
2	×	×	×	×	×	×	×
3	×	×	×	×	×	×	×
4	×	×	×	×	×	×	×

The experiment: 3 blocks, each with 35 jars

Environment Complexity	Assemblage identity						
	A	B	C	AB	AC	BC	ABC
0	×	×	×	×	×	×	×
1	×	×	×	×	×	×	×
2	×	×	×	×	×	×	×
3	×	×	×	×	×	×	×
4	×	×	×	×	×	×	×

Spanish PhD student Lorea Flores visited the University of Roehampton for three months;
gathered the “shrimps” from ponds on the campus;
put the combinations of leaves, shrimps and plastic rings into jars;
put one jar of each type onto each of three shelves in a temperature-controlled room;
measured various responses on each jar (some daily, some at the end).

Models and data analysis

The models consist of all interactions and sums of those shown in the two previous diagrams
(the gentle reader can draw her own Hasse diagram!).

Models and data analysis

The models consist of all interactions and sums of those shown in the two previous diagrams
(the gentle reader can draw her own Hasse diagram!).

Analysis of variance is the standard statistical technique which enables us to find the most parsimonious model which explains the data adequately.

Models and data analysis

The models consist of all interactions and sums of those shown in the two previous diagrams
(the gentle reader can draw her own Hasse diagram!).

Analysis of variance is the standard statistical technique which enables us to find the most parsimonious model which explains the data adequately.

RAB gets the data sheet, works out how to do the analysis, and simply gets out her hand calculator . . .

Models and data analysis

The models consist of all interactions and sums of those shown in the two previous diagrams
(the gentle reader can draw her own Hasse diagram!).

Analysis of variance is the standard statistical technique
which enables us to find the most parsimonious model
which explains the data adequately.

RAB gets the data sheet, works out how to do the analysis,
and simply gets out her hand calculator …

… but the ecologists cannot do this. They can use statistical software to fit each model, and then use a spreadsheet to subtract sums of squares appropriately.
This is error-prone.

Models and data analysis

The models consist of all interactions and sums of those shown in the two previous diagrams
(the gentle reader can draw her own Hasse diagram!).

Analysis of variance is the standard statistical technique which enables us to find the most parsimonious model which explains the data adequately.

RAB gets the data sheet, works out how to do the analysis, and simply gets out her hand calculator …

… but the ecologists cannot do this. They can use statistical software to fit each model, and then use a spreadsheet to subtract sums of squares appropriately. This is error-prone.

Solution! Summer student Justin Thong dug into the statistical software R to find a short sequence of commands that gives precisely the right output
(not straightforward, because R makes some stupid assumptions).

Hooray!

PLOS | ONE

RESEARCH ARTICLE

Habitat Complexity in Aquatic Microcosms Affects Processes Driven by Detritivores

Lorea Flores[1]*, R. A. Bailey[2,3], Arturo Elosegi[4], Aitor Larrañaga[4], Julia Reiss[5]*

1 INRA, UMR 1224, Ecologie Comportementale et Biologie des Populations de Poissons, Aquapôle, quartier Ibarron, 64310 Saint-Pée sur Nivelle, France, **2** School of Mathematical Sciences, Queen Mary University of London, London E1 4 NS, United Kingdom, **3** School of Mathematics and Statistics, University of St Andrews, St Andrews, Fife, KY16 9SS, United Kingdom, **4** Laboratory of Stream Ecology, Dept. of Plant Biology and Ecology, Fac. of Science and Technology, University of the Basque Country, UPV/EHU PO Box 644; 48080 Bilbao, Spain, **5** Department of Life Sciences, Whitelands College, University of Roehampton, London SW15 4JD, United Kingdom

* lorea.flores@st-pee.inra.fr (LF); julia.reiss@roehampton.ac.uk (JR)

CrossMark
click for updates

Abstract

So what affected the three measured responses?

Individual species numbers;
Plastic rings or not;
Number of plastic rings.

Nothing more complicated, so
not Richness,
not Fractal dimension,
no interactions.