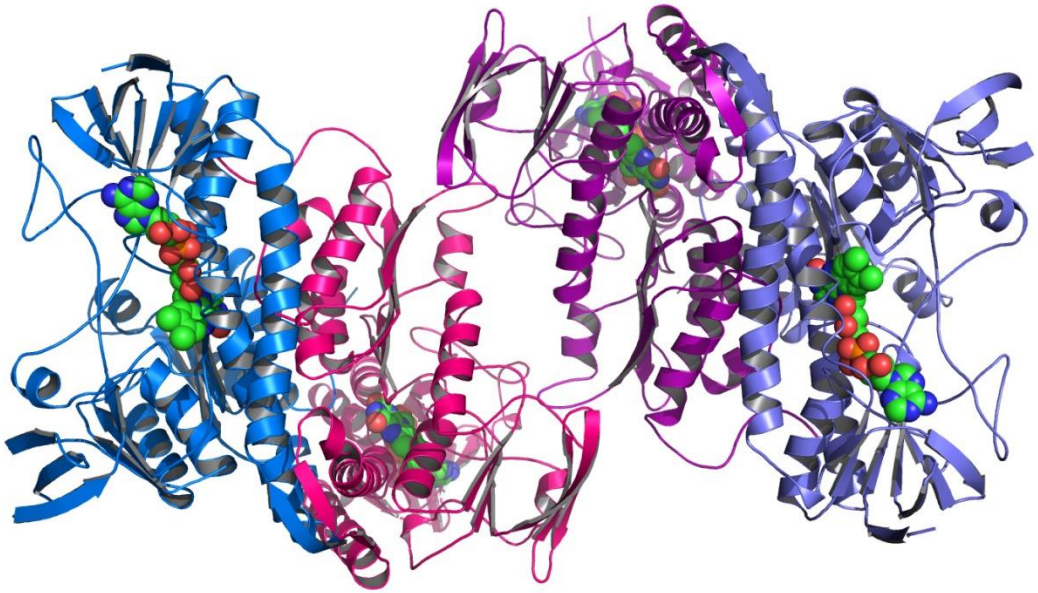
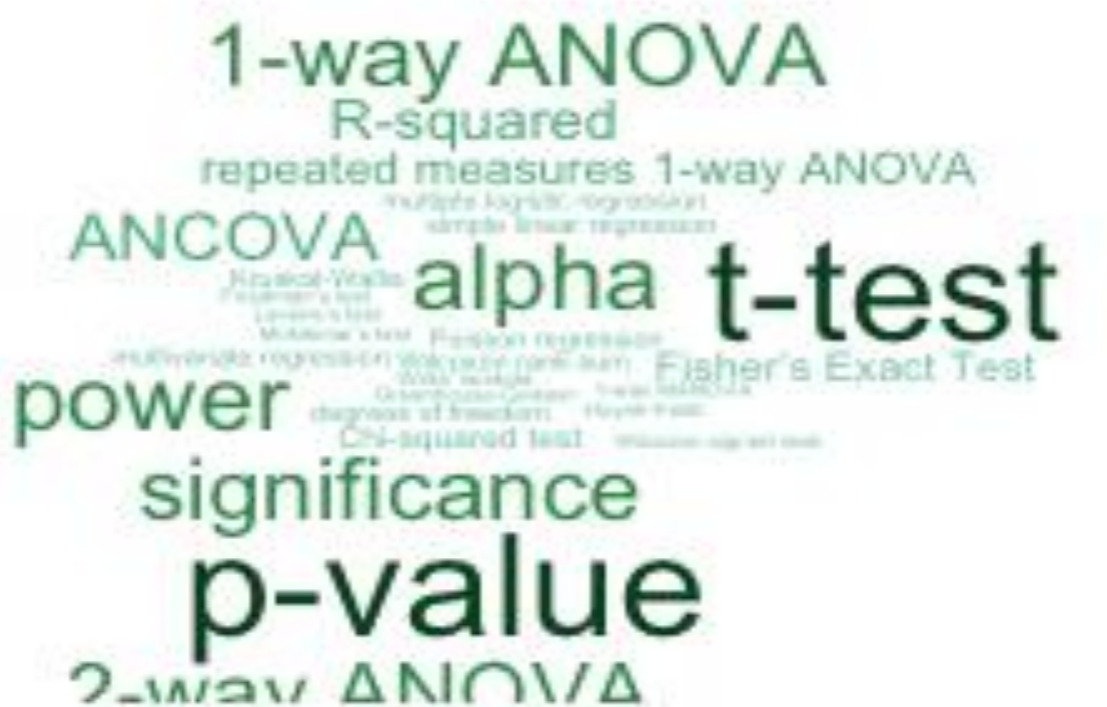
BIOL 8700: Experimental Plans
Part 1

Terry Neeman, PhD AStat
Biology Data Science Institute (BDSI)
Joint Colleges of Science, ANU



What a first course in statistics emphasises….



What statistics is really about:

The centrepiece of statistics: Models, not p-values

Model: a descriptive summary of the data

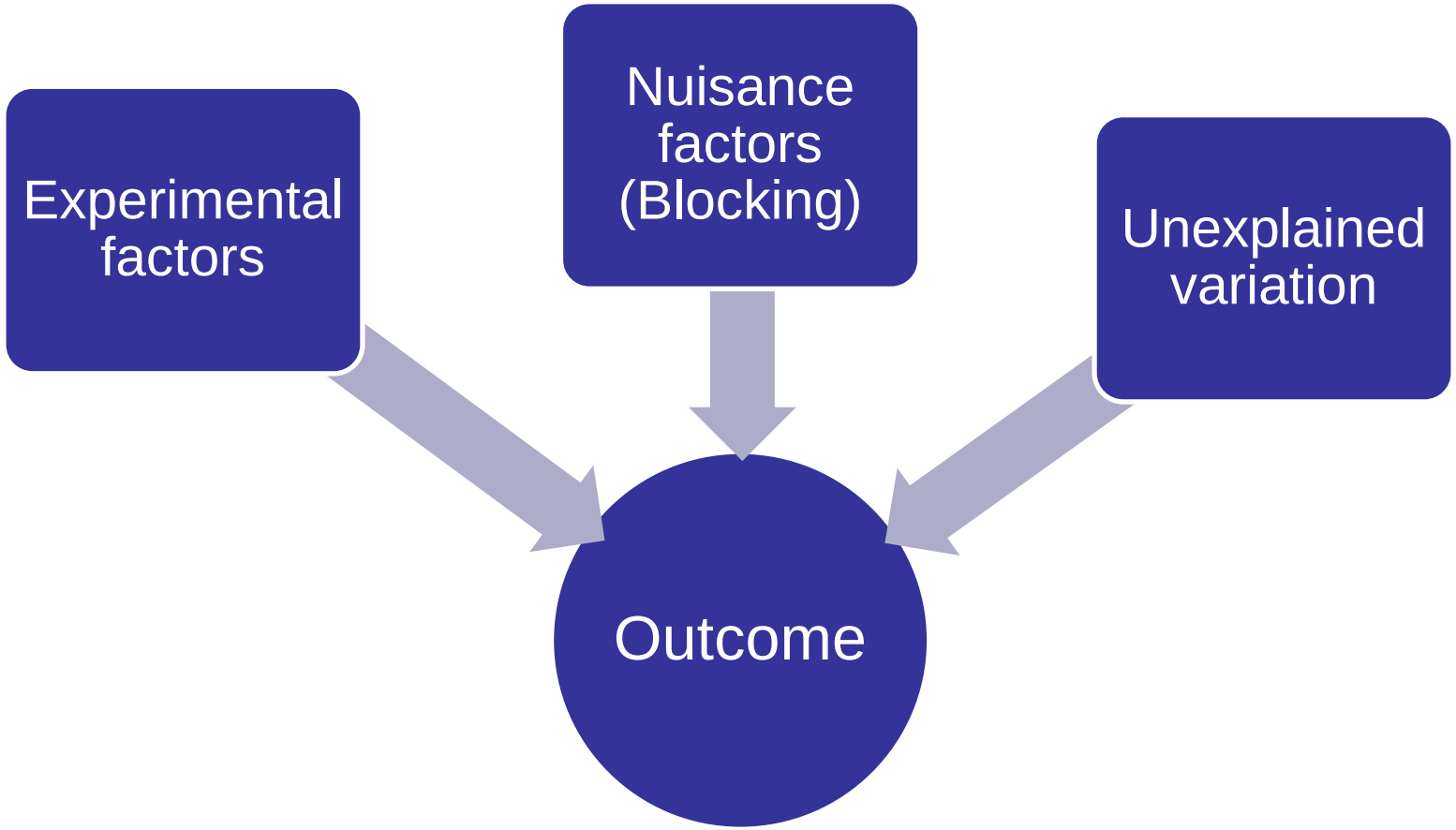
Model: telling the “big picture” story of the data

Model: an explanatory summary of the data

Model: a mechanism for predicting outcomes

Model: turning data into information.

Model of an Experiment

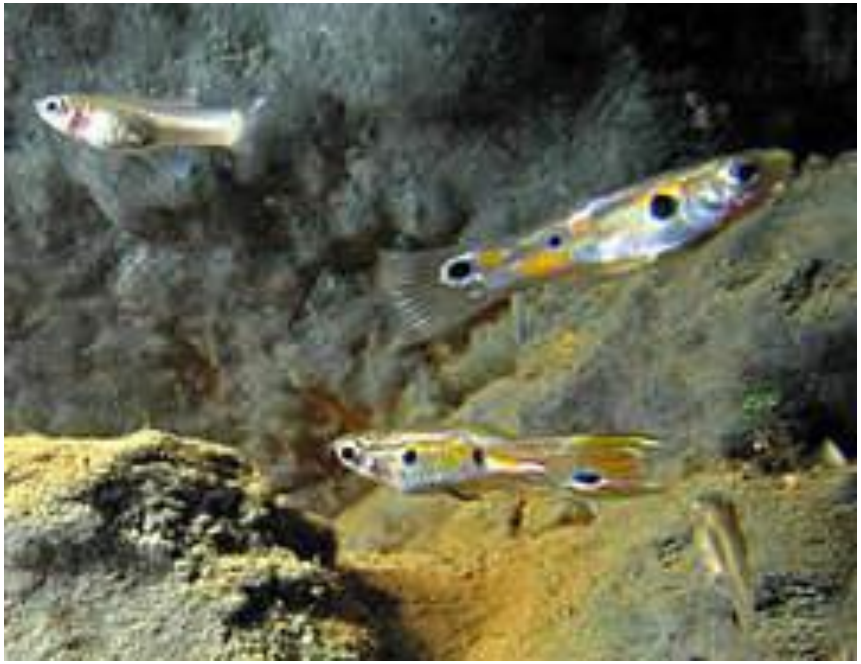


When you just think about doing some statistical test, you may forget about other important sources of variation….

What is effect of LPS challenge on sexually selected traits in guppies?

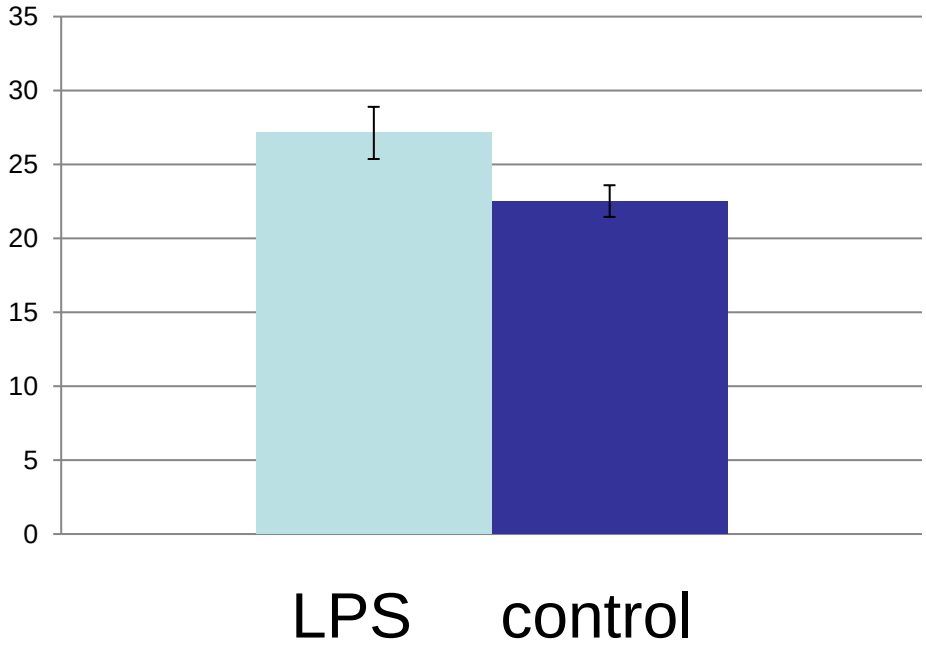
Tank 1: control guppies

Tank 2: LPS-injected guppies

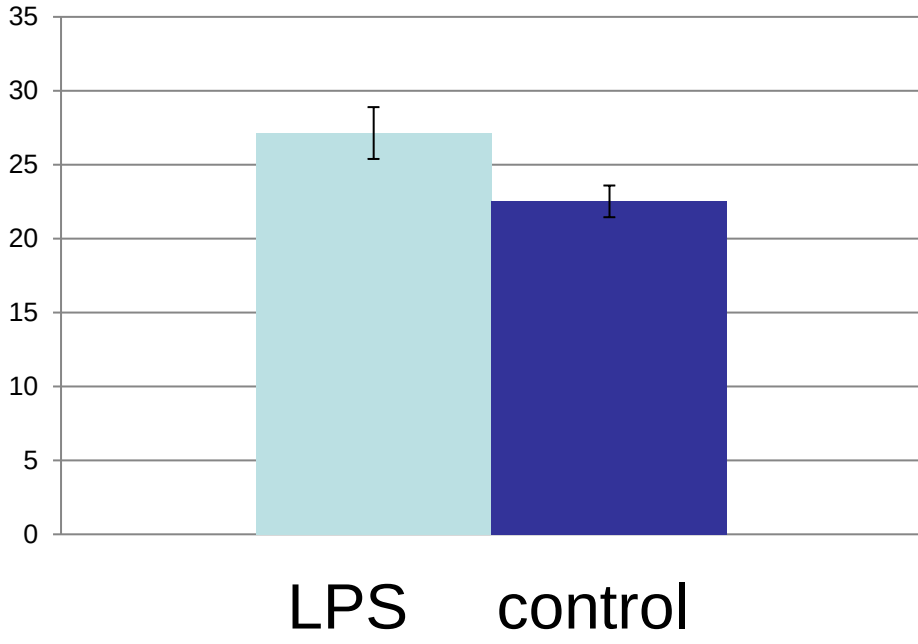


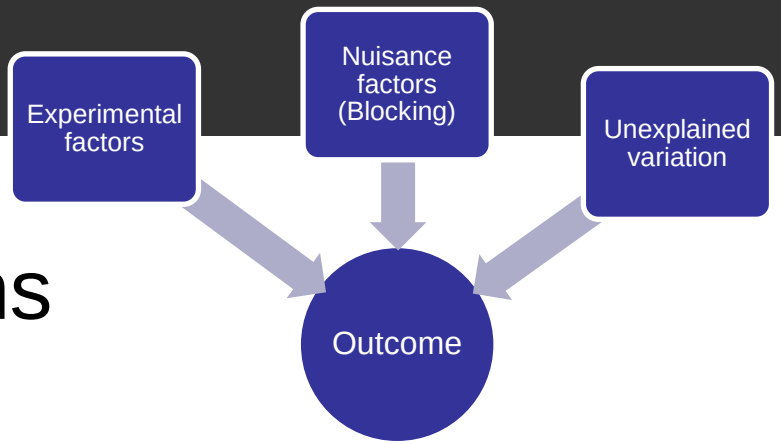
Compare mean sperm velocity pre- and post-

Pre



Post





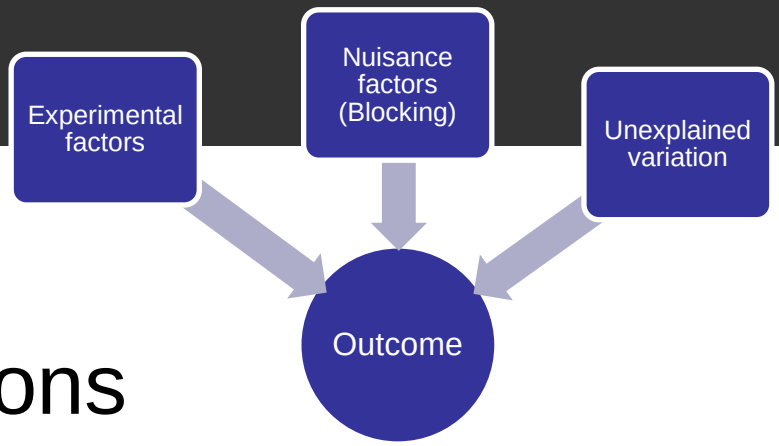
Questions

Nuisance factors:

Are the differences observed due to Treatment or Tank?

Could the differences observed be due to differences in fish size between groups?

What are the potential nuisance factors in this experiment?



More Questions

Experimental factors:

Does LPS injection impact response?

Does the LPS effect different post vs pre?

Is there an overall time effect (post vs pre)?

How can we ensure that the LPS effect is not confounded with the *nuisance factors*?

Model-centered thinking Principles of Experimental Design

KEY PRINCIPLES in Experimental Design and Analysis

- CONTROLS
- REPLICATION
- BLOCKING
- RANDOMISATION
- BLINDING

KEY PRINCIPLES in Experimental Design and Analysis

- CONTROLS:
  - Direct comparison with a known standard or no treatment.
  - Tested under identical conditions to experimental treatment.
- REPLICATION
- BLOCKING
- RANDOMISATION
- BLINDING

KEY PRINCIPLES in Experimental Design and Analysis

- CONTROLS
  - Direct comparison with a known standard or no treatment.
  - Tested under identical conditions to experimental treatment.
- REPLICATION: repeating experiment on different samples to:
  - (1) increase precision of treatment effect
  - (2) make result more generalisable
- BLOCKING
- RANDOMISATION
- BLINDING

KEY PRINCIPLES in Experimental Design and Analysis

CONTROLS

- Direct comparison with a known standard or no treatment.
- Tested under identical conditions to experimental treatment.

REPLICATION: repeating experiment on different samples to:

(1) increase precision of treatment effect

(2) make result more generalisable

- BLOCKING:
  - Grouping together similar experimental units, &
  - Comparing treatments within homogeneous groups
- RANDOMISATION
- BLINDING

KEY PRINCIPLES in Experimental Design and Analysis

•CONTROLS

- Direct comparison with a known standard or no treatment.
- Tested under identical conditions to experimental treatment.

•REPLICATION repeating experiment on different samples to:

(1) increase precision of treatment effect

(2) make result more generalizable

•BLOCKING

- Grouping together similar experimental units, &
- Comparing treatments within homogeneous groups

•RANDOMISATION: probabilistic process of assigning treatment.

•BLINDING

KEY PRINCIPLES in Experimental Design and Analysis

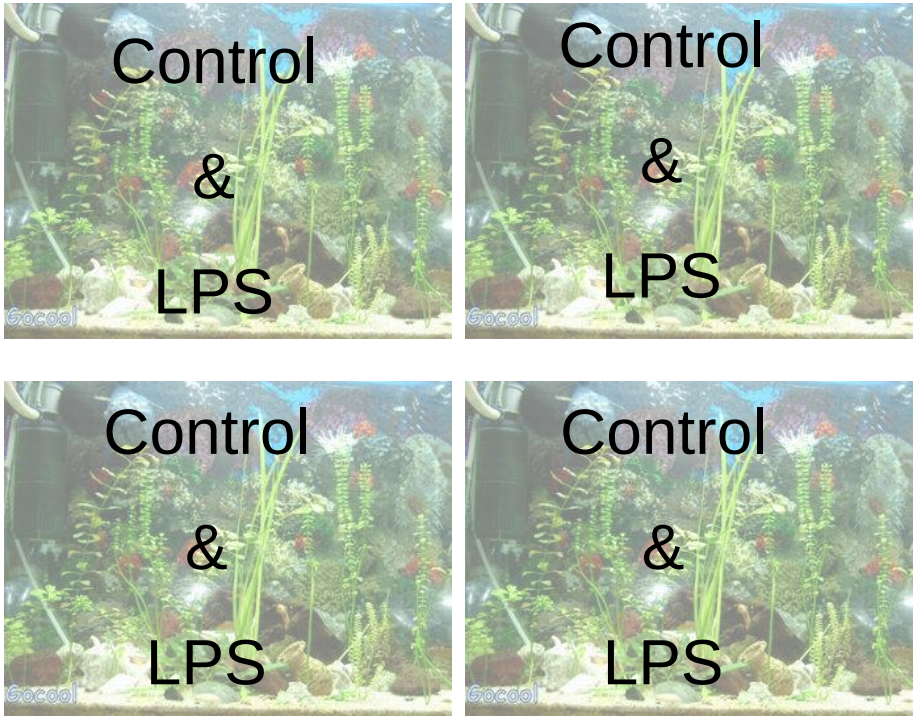
- CONTROLS
  - Direct comparison with a known standard or no treatment.
  - Tested under identical conditions to experimental treatment.
- REPLICATION repeating experiment on different samples to:
  - (1) increase precision of treatment effect
  - (2) make result more generalisable
- BLOCKING
  - Grouping together similar experimental units, &
  - Comparing treatments within homogeneous groups
- RANDOMISATION probabilistic process of making treatment assignments.
- BLINDING: masking treatment assignment.

Let’s improve our experimental design in the guppy experiment

More tanks?

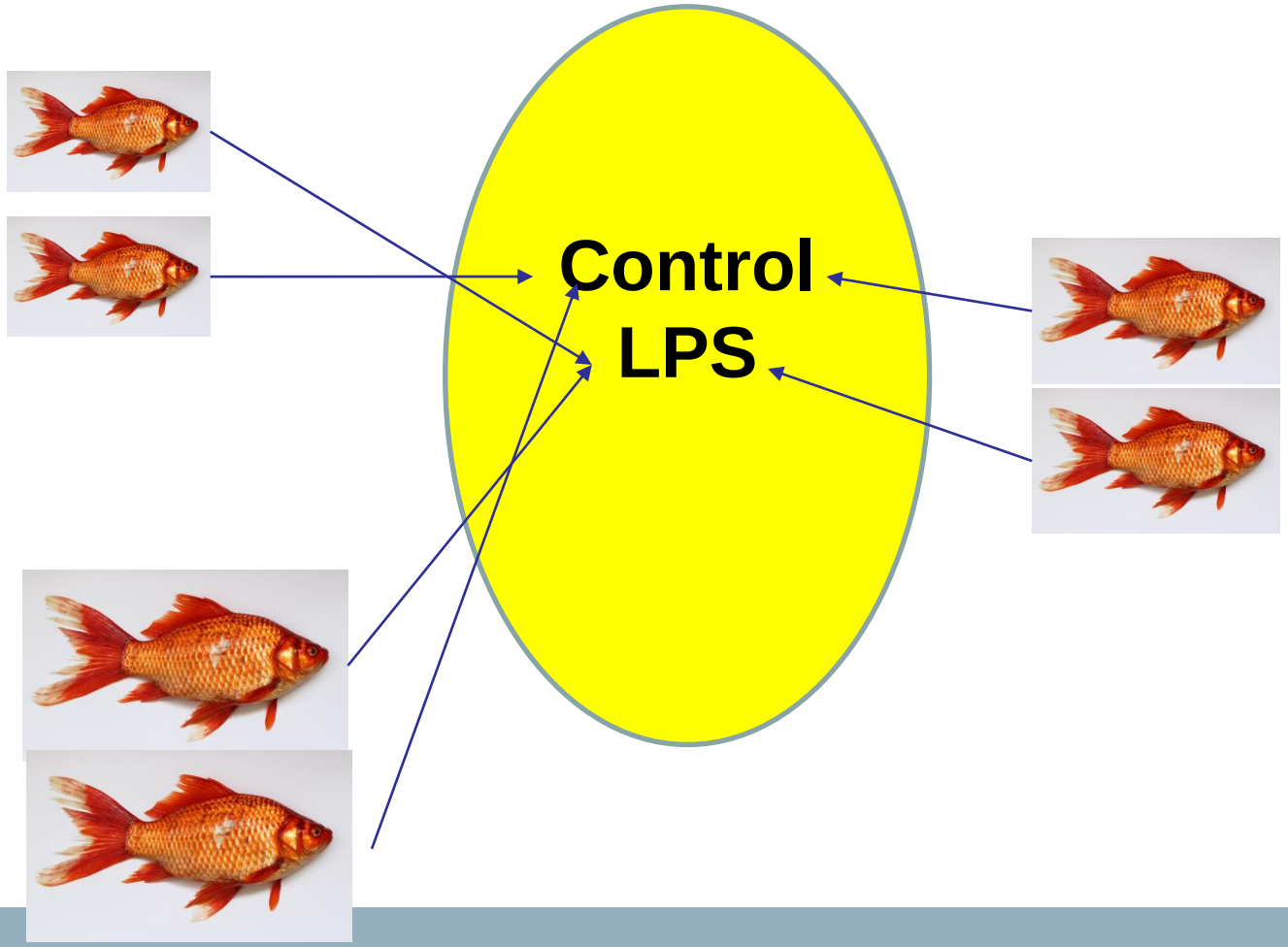
Let’s improve our experimental design in the guppy experiment

Mixed tanks?



Let’s improve our experimental design

Randomise to treatment stratify by size



How else can we ensure that our control group is a good comparison for the LPS group?

- Mock injections?
- Timing of measurements?
- Evaluator-blinded?

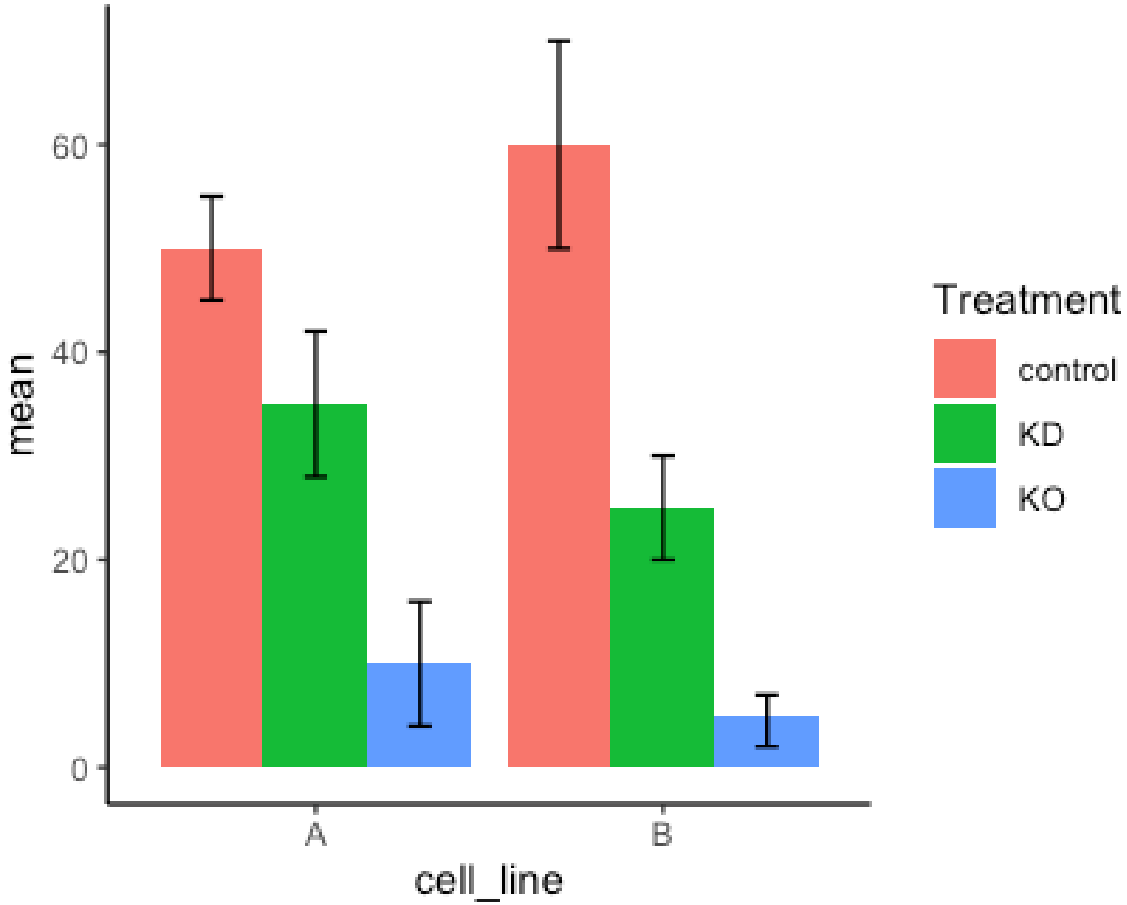
Summary so far

- Goal of experimental design: separate out effects due to treatment from effects from (nuisance) experimental factors.
- Good experimental design maximises the information about treatment effect.
- Model-centred thinking informs the design of experiments.
- Models incorporate blocking factors into the analysis of treatment effects.

Experiment 2:

Effect of LCP1 knock-out/knock-down on cytoskeleton structure in cell lines

Compare mean outcome on cytoskeleton



What is effect of LCP1 KO/KD on cytoskeleton structure in cell lines A and B?

Experimental factor:

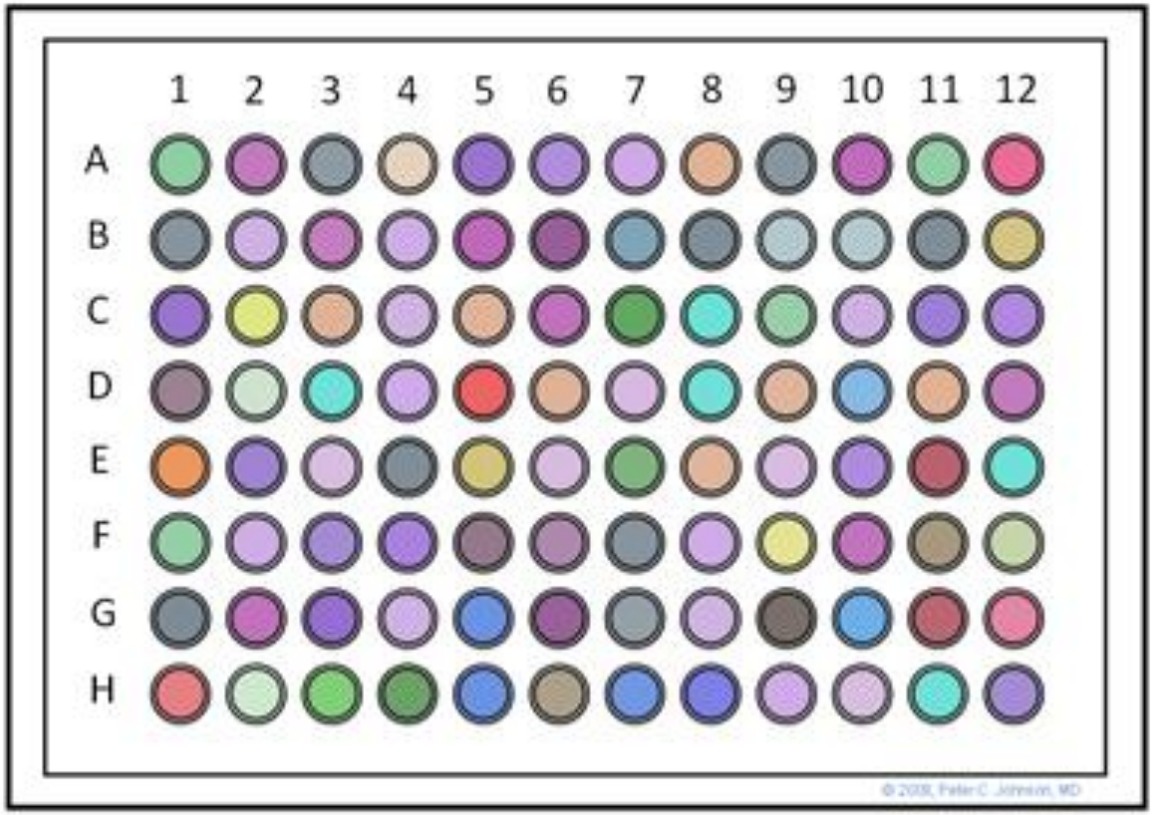
LCP1: present/KD/KO

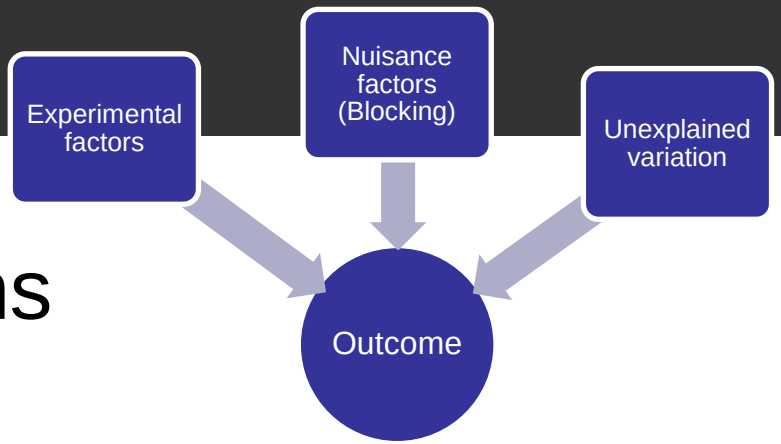
Blocking factors:

Cell line (A/B)

Plate

Position on plate





Questions

Nuisance factors:

Are the differences observed due to Plate or Position/Plate?

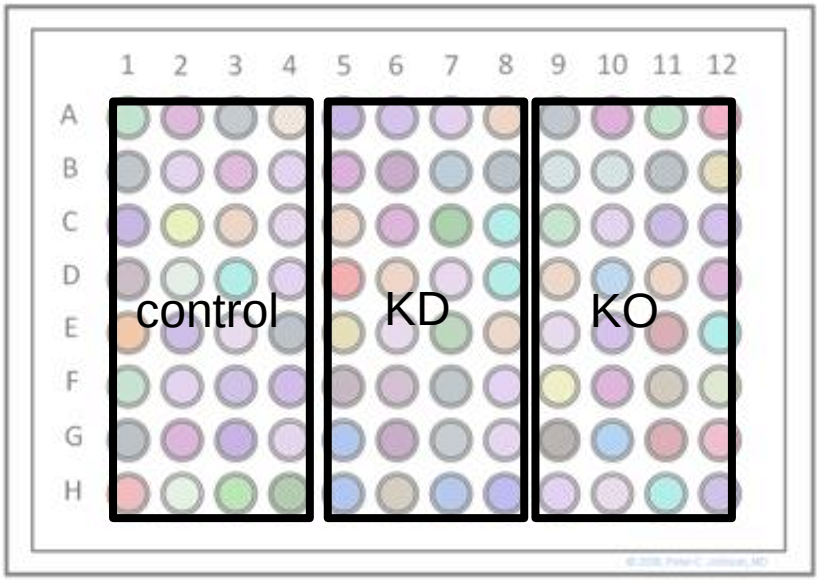
How many plates will you use?

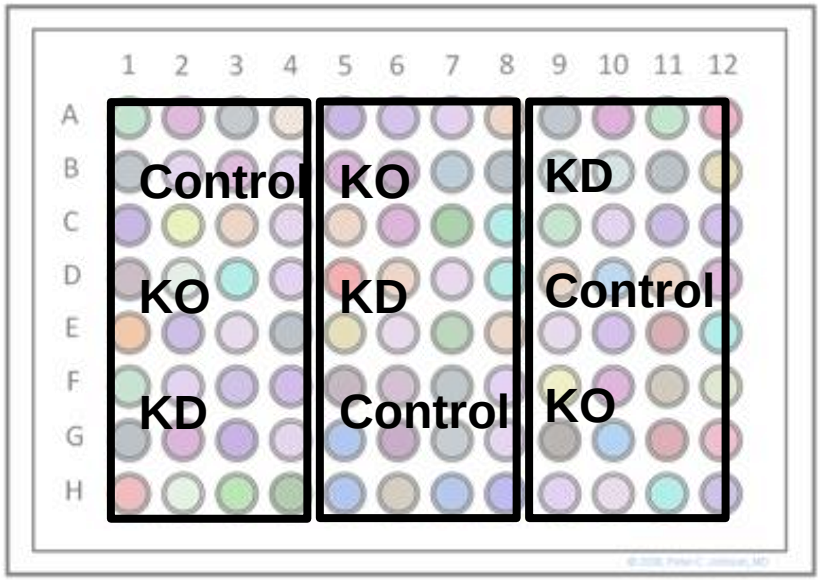
Will there be multiple wells per treatment/cell line?

If so, are these biological replicates or technical replicates?

One plate per cell line, or both cell lines on each plate?

How would you design your experiment?

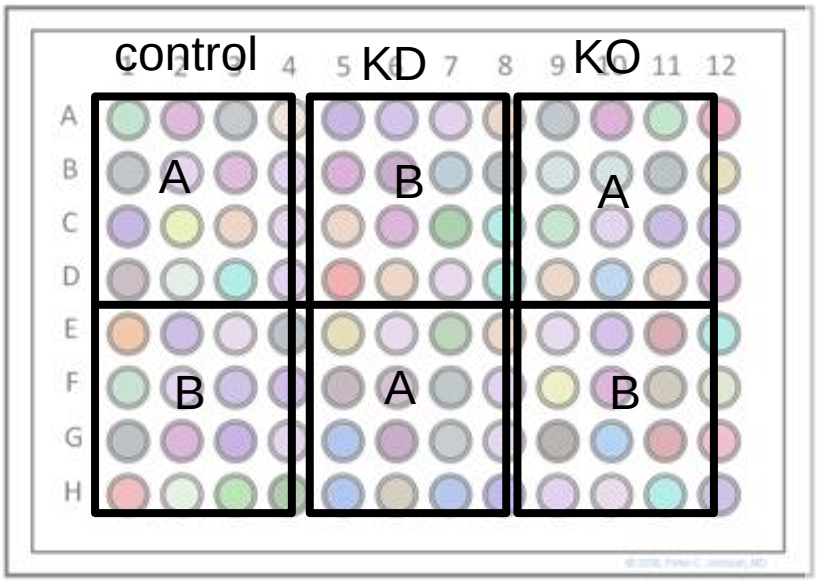


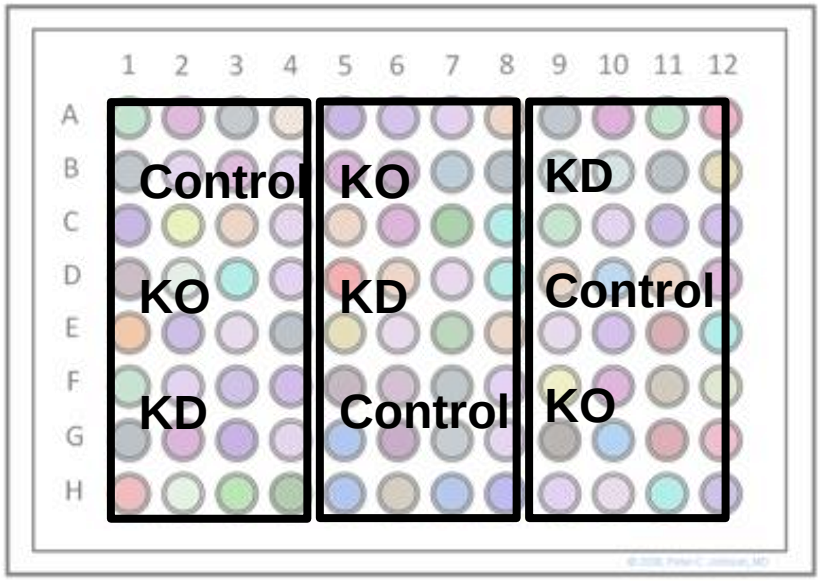


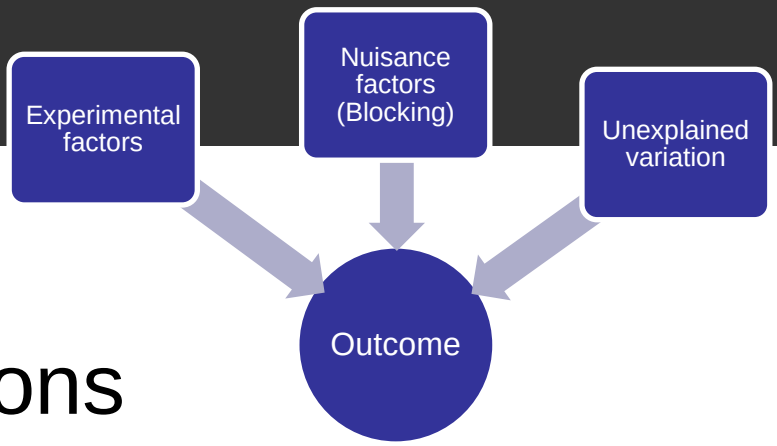
How would you arrange by cell line?

Good or bad design?

Plate 1: A. Plate 2: B







More Questions

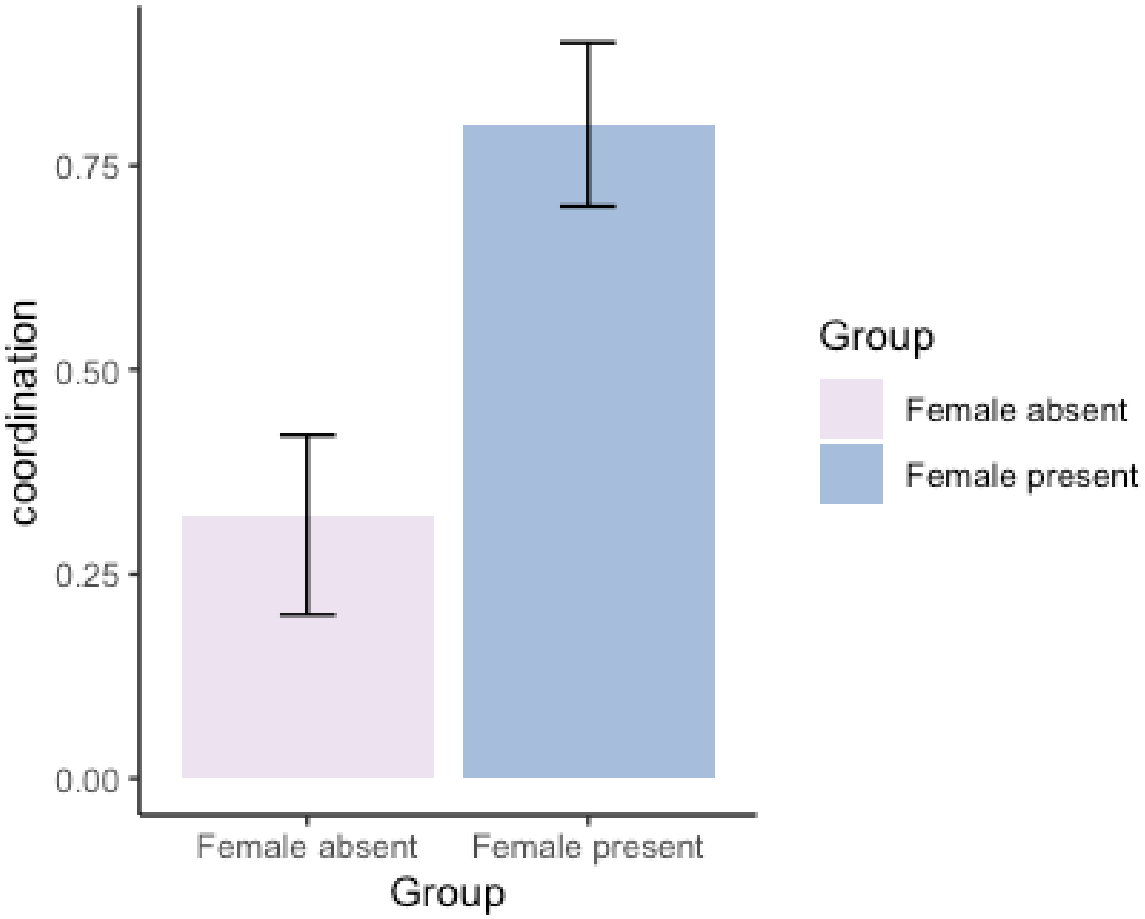
Experimental factor:

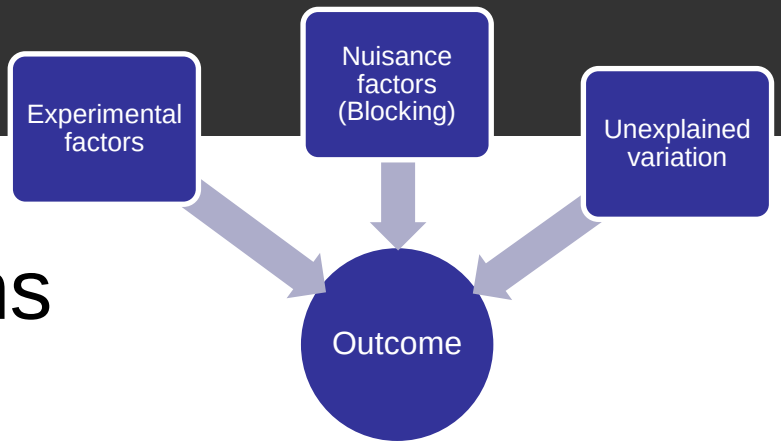
Does Treatment impact response?

How can we ensure that Treatment is not confounded with the *nuisance factors*?

Experiment 3:
Differences in synchronous waving patterns in fiddler crabs when females present/absent

Compare mean outcome





Questions

Nuisance factors:

Are the differences observed due to Cluster?

How many clusters will you use?

Could each cluster receive both treatments?

What are the advantages of assigning both treatments to a cluster?

If so, could the order of treatments matter?

How would you design your experiment?

Cluster 1

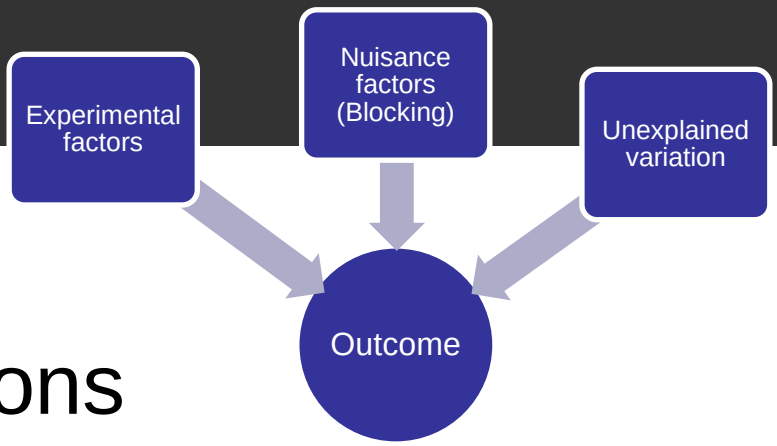
Cluster 3

Female
- Absent
- Present

Cluster 2

How many runs within each cluster?

Cluster 4



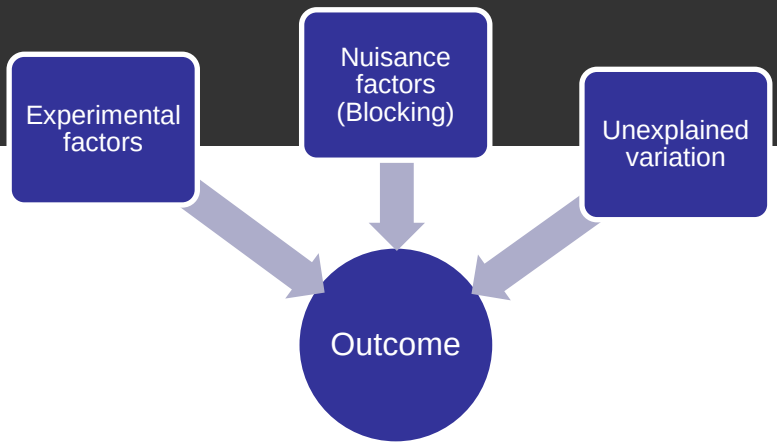
More Questions

Experimental factor:

Does female presence/absence impact response?

How can we ensure that the female effect is not confounded with the *nuisance factors*?

Observational Study:
Fish community structure on the Ningaloo reef



Independent experimental variables

Patch size

Patch shape

Perimeter: area ratio

Edge length

Position (i.e. edge vs interior)

Dependent experimental variables

Fish community structure

•• Species Richness

•• Diversity

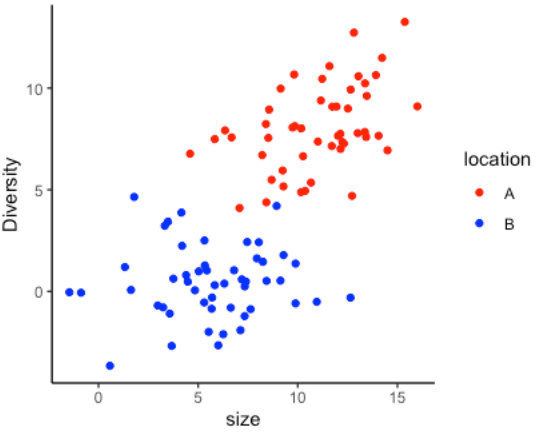
•• Relative abundance

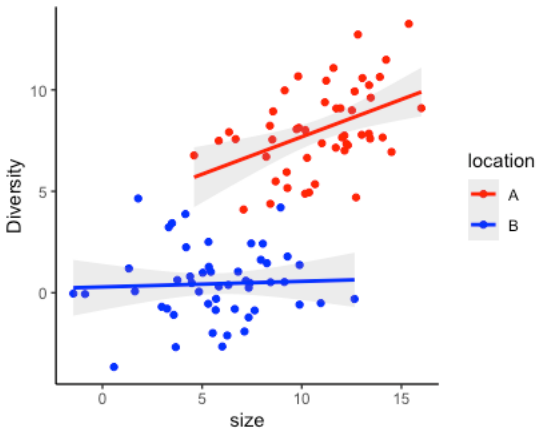
•• Functional trophic groups present

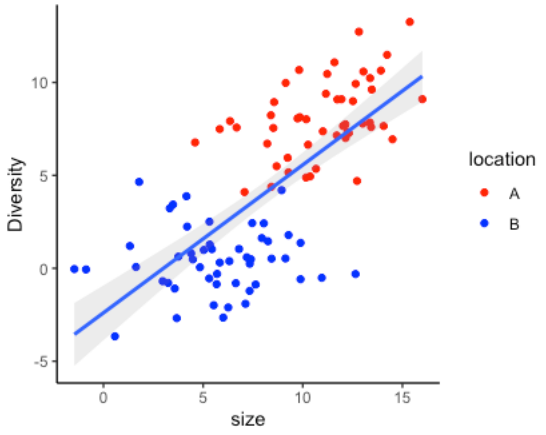
Experimental unit = patch

Location can confound the effect of patch size on diversity

How can good survey design be utilized?







Other points to consider

1. Think about how location within the reef can impact on response. Treat location as “block”.

2. Patches within each location as different as possible with respect to size, shape complexity, edge length etc.
3. How many locations? How many patches within location?
4. Position effects are assessed WITHIN each patch
5. Size effects are assessed BETWEEN patches.

Model-centered thinking: R, modelling and reproducible research

R: your friendly data analysis tool

- Advantages:
  - Great graphics tools for data visualization and pattern recognition
  - Easy to model data, interpret and summarise data
  - Helps with data/project organization
  - Essential tool for reproducible research/ transparency in research
- Disadvantages:

R: your friendly data analysis tool

- More Advantages:
  - Other researchers can follow logic of analysis
  - Can make code/data available for future researchers
  - Large community of R users internationally
    - Use latest analytic techniques (R packages)
    - Ask questions to experts in field on good analyses
- Disadvantages:

R: your friendly data analysis tool

- More Advantages:
  - Code compiles into beautifully formatted document
  - R coding skills valued inside/outside academia
- Disadvantages:
  - Can be a steep learning curve
  - Can be hard to get started/ not enough local resources

Example of analysis using R

- General workflow
  - Import data
  - General data management
  - Exploratory data analysis
  - Fit model to data
  - Check model assumptions, interpret model output
  - Summarise model using graphics

Example of analysis using R (markdown):
Analysis of root growth.docx