Experimental Design
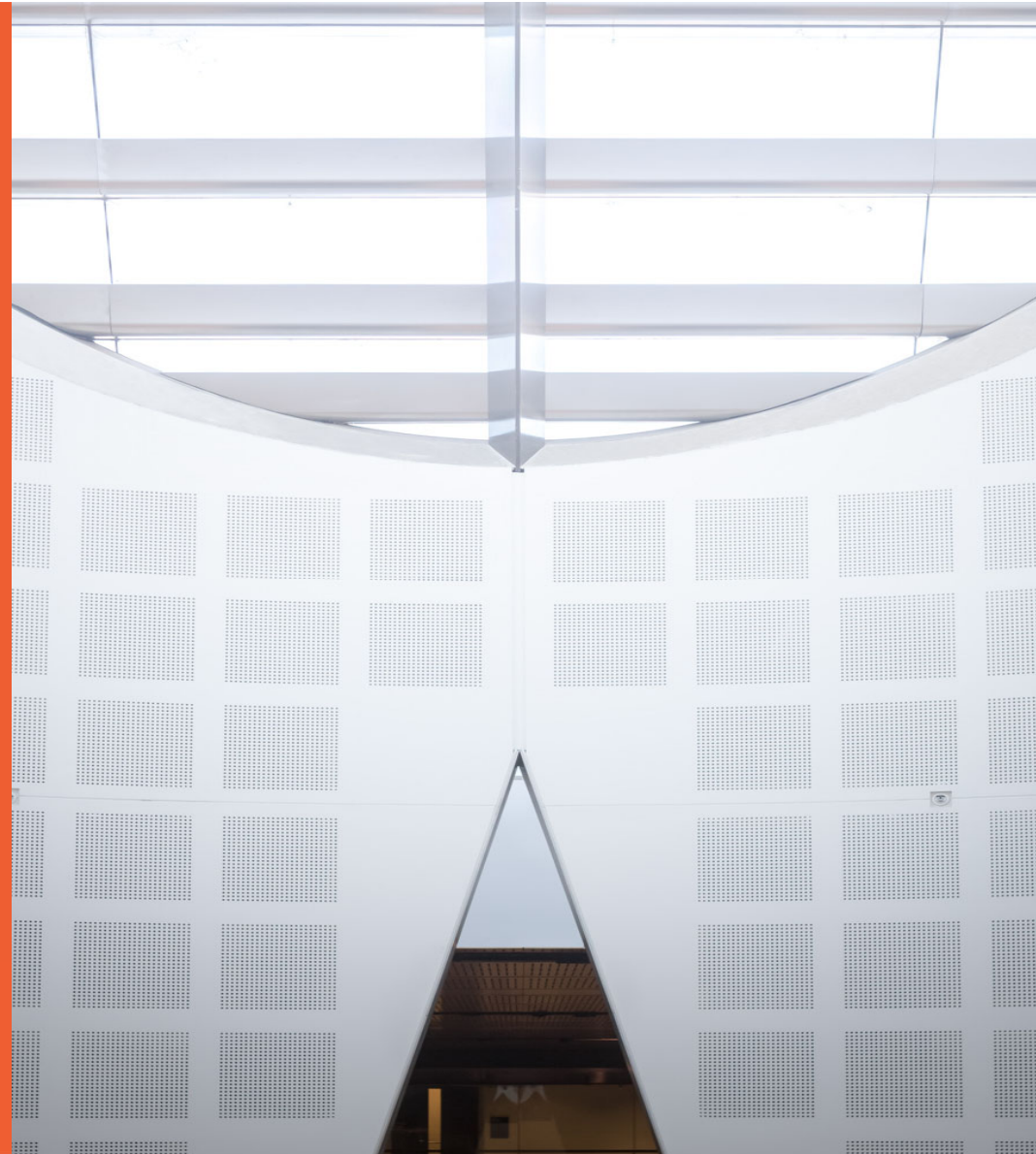
Presented by
Alex Shaw
Statistical Consultant
Sydney Informatics Hub
Core Research Facilities
The University of Sydney

THE UNIVERSITY OF
SYDNEY

Acknowledging SIH

All University of Sydney resources are available to Sydney researchers **free of charge.** The use of the SIH services including the Artemis HPC and associated support and training warrants acknowledgement in any publications, conference proceedings or posters describing work facilitated by these services.

The continued acknowledgment of the use of SIH facilities ensures the sustainability of our services.

Suggested wording:

General acknowledgement:

"The authors acknowledge the technical assistance provided by the Sydney Informatics Hub, a Core Research Facility of the University of Sydney."

Acknowledging specific staff:

"The authors acknowledge the technical assistance of (name of staff) of the Sydney Informatics Hub, a Core Research Facility of the University of Sydney."

For further information about acknowledging the Sydney Informatics Hub, please contact us at **sih.info@sydney.edu.au**.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

How to use this workshop

Workshops developed by the Statistical Consulting Team within the Sydney Informatics Hub form an integrated modular framework. Researchers are encouraged to choose modules to *create custom programmes tailored to their specific needs.* This is achieved through:

- Short 90 minute workshops, acknowledging researchers rarely have time for long multi day workshops.

- Providing statistical workflows appliable in any software, that give practical step by step instructions which researchers return to when analysing and interpreting their data or designing their study e.g. workflows for designing studies for strong causal inference, model diagnostics, interpretation and presentation of results.

- Each one focusing on a specific statistical method while also integrating and referencing the others to give a holistic understanding of how data can be transformed into knowledge from a statistical perspective from hypothesis generation to publication.

For other workshops that fit into this integrated framework refer to our training link page under statistics https://www.sydney.edu.au/research/facilities/sydney-informatics-hub/workshops-and-training.html#stats

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

What is a workflow?

– Every statistical analysis is different, but all follow similar paths. It can be useful to know what these paths are.

– We have developed practical, step-by-step instructions that we call '*workflows*', that can you can follow and apply to your research

– We have a general research workflow that you can follow from hypothesis generation to publication

– And statistical workflows that focus on each major step along the way (e.g. experimental design, power calculation, model building, analysis using linear models/survival/multivariate/survey methods)

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Statistical Workflows

- Our statistical workflows can be found within our workshop slides.

- Statistical workflows are software agnostic, in that they can be applied using any statistical software.

- There may also be accompanying software workflows that show you how to perform the statistical workflow using particular software packages (e.g. R or SPSS). We won't be going through these in detail during the workshop. If you are having trouble using them, we suggest you attend our monthly Hacky Hour where SIH staff can help you.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

During the workshop

Ask short questions or clarifications during the workshop. There will be breaks during the workshop for longer questions.

Slides with this blackboard icon are mainly for your reference, and the material will not be discussed during the workshop.

Challenge Question

– A wild boar is coming towards you at 20mph. Do you:?
  – A. Ask it directions
  – B. Wave a red flag
  – C. Wave a white flag
  – D. Begin preparing a trap

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

General Research Workflow

1. **Hypothesis Generation** (Research/Desktop Review)
2. <mark>**Experimental and Analytical Design** (sampling, power, ethics approval)</mark>
3. **Collect/Store Data**
4. **Data cleaning**
5. **Exploratory Data Analysis (EDA)**
6. **Data Analysis aka inferential analysis**
7. **Predictive modelling**
8. **Publication**

This workshop

I.   Statistical principles underpinning experimental design

II.   A workflow for experimental design

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Statistical Principles Underpinning Experimental Design

Why do we care about experimental design?

The big picture - sampling

– Ronald Reagan loved jellybeans

– If we wanted to estimate the proportion of red jellybeans in this jar, we could count them all, or take a **sample**

– Taking a sample, and using it to make **estimates** and draw **inferences** about a **population** (in this case all of the jelly beans in the jar) is the basis of statistics.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

# Estimates are inaccurate and imprecise… but by how much?

– When we take a sample, and run a study there are two 'forces' that make our estimates differ from the true value in the population: **bias** (accuracy) and **error** (precision)

– Error and bias are quite different. Bias is unknown, whereas error is in some sense more 'visible' because it can at least be estimated from a single experiment

– Experimental design is all about about controlling **bias** and minimising **error** (variability) in order to estimate the true value as closely as possible

| Accurate Precise | Not Accurate Precise | Accurate Not Precise | Not Accurate Not Precise |
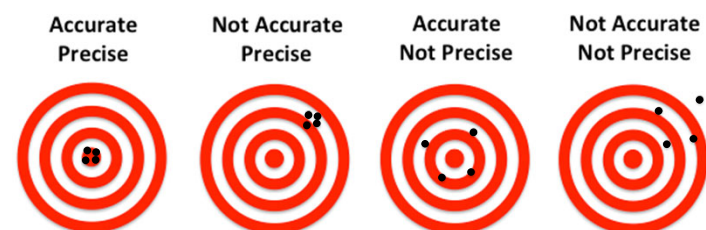
Just like throwing a dart, using our experimental sample to estimate some population statistic is subject to randomness. Unlike throwing a dart – we don't know where the bullseye is!

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Why are our estimates biased?

- Let's start by thinking about bias in a statistical way: it is *systematically* over- or under-estimating our population property

- Bias may be introduced in the way we **sample, conduct our experiment, measure,** or even **analyse** our data.

- Some level of bias is inevitable, but too much bias can be fatal: we would reach the wrong conclusion *most of the time* in a poorly designed experiment!
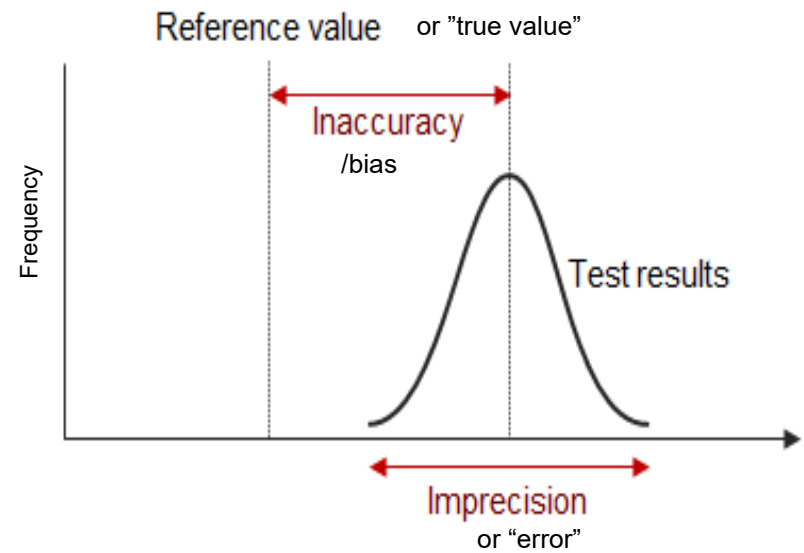
If we ran our experiment many times, and plotted the frequency of our estimates, our best guess of the true value would be the mean of all of our estimates. The distance from this best guess to the true value is our bias. No matter how many times we ran the experiment, we would end up with this level of bias.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Why are our estimates imprecise?

- 'Error' in statistical terms is about how close together our estimates are — it's 'error' in the sense of measurement, not making a mistake

- Error exists because the individuals in our population vary

- No matter how large our sample, our estimates will vary to some extent, which is the same reason we use confidence intervals and p-values to report our estimates

- Again, error is influenced by the way we sample, measure, conduct or analyse our experiment

- Choosing an appropriate sample size is a vital and non-trivial part of controlling error in experimental design: see our Power and Sample Size Calculation workshop for more
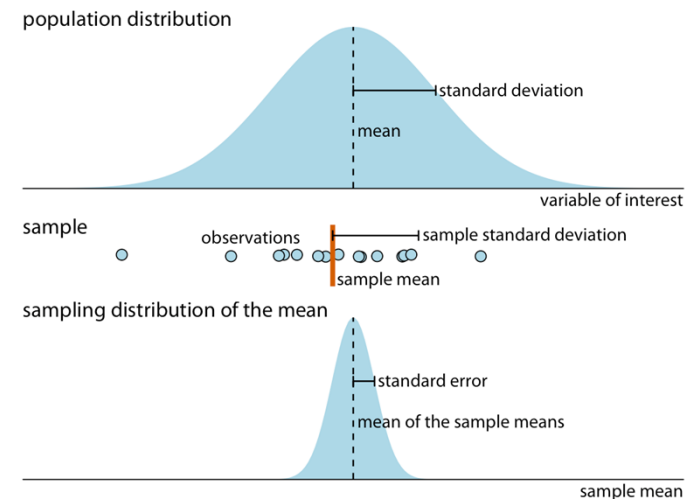
It's important to realise that although variability in the population affects the variability of your estimates –there's another important ingredient: your sample size.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Questions so far?

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

A quick challenge question

– Would you be concerned about bias, error, or both in the following studies?

- – A diet study in which respondents were asked to recall their food intake for the previous week every Sunday evening

- – An animal behaviour study that recruited subjects by asking owners to bring in their pet dogs of any breed

- – A survey on the number of motor vehicle crashes experienced over your life time where you, the respondent was at fault

- – A study that tested the nutritional effects of milk on children, recruited through a school newsletter asking parents to volunteer their children

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

A workflow for experimental design
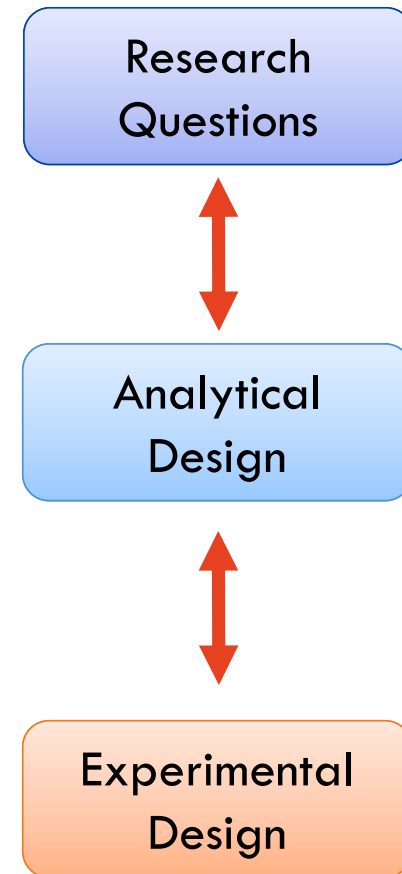
Experimental Design Workflow

Step 0) Finalise your research questions and hypotheses

Step 1) Identify your study type, target population, sampling frame

Step 2) Consider analysis of your data in chosen study design including any treatments and define units

Step 3) Design your study considering:
- Control of bias including randomisation and blinding
- Variability and replication: blocking, power calculation to assess needed replication

Research Questions

Analytical Design

Experimental Design

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Step 0) Finalise your research question and hypotheses

What are your research questions? Having a very clear idea of what your research questions is vital for achieving good experimental design.

Avoid the urge to rush into collecting data and just "worry about the analysis later."

You can only get meaningful results from a well-controlled and adequately-powered experiment. The type of analyses possible are also dictated by your experimental design.

Think of this paradox: *often the sooner you start your experiments, the longer your project will take.* Invest the time in good design.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

1. Identify your study type, target population, sampling frame

0. Formulate your research question

– What are your research questions?

– What is the best approach to answer your research question?

– Many approaches are available:
  – Field experiment
  – Natural experiment
  – Laboratory experiment
  – Systematic review and meta-analysis
  – Randomised controlled trial
  – Crossover study
  – Clinical study
  – Case-control study
  – Cross-sectional study
  – Longitudinal study
  – Interrupted time series design
  – Other???

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney
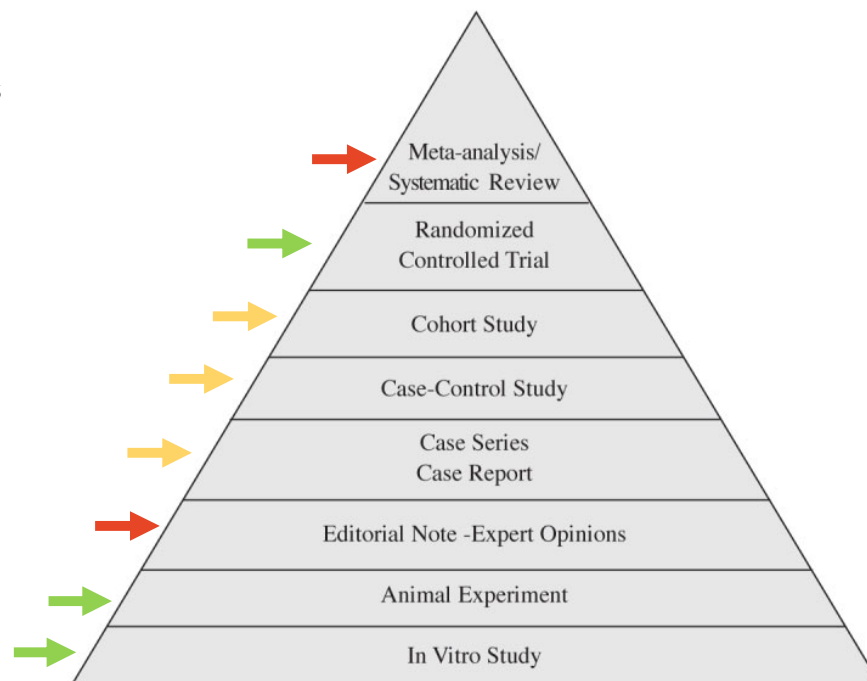
1. Identify your study type

Is there a best study type?

- Consider this evidence pyramid.

- Some will involve be designed experiments and others are observational, while others involve synthesis of primary literature

- The optimal study type will depend on what is known and the feasibility of each study type in your chosen field

- You can't proceed to an RCT without laying the base of the pyramid, and even then an RCT may not be 'best' to study type for your phenomena of interest

Study designs in medicine
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4318396/

The evidence pyramid and introduction to randomized controlled trials. *Pandis N Am J Orthod Dentofacial Orthop. 2011 Sep; 140(3):446-7.*

Evidence pyramid for medical studies

Meta-analysis/ Systematic Review
Randomized Controlled Trial
Cohort Study
Case-Control Study
Case Series Case Report
Editorial Note -Expert Opinions
Animal Experiment
In Vitro Study

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

1. Observational vs. Designed Experiments

- In studying some phenomena in the real world we are usually interested in **causality.** E.g. In health research we may aim to find modifiable factors that can improve health outcomes.

- **A causal** relationship between two variables refers to an explanatory variable affecting an outcome variable (e.g. diet influencing weight)
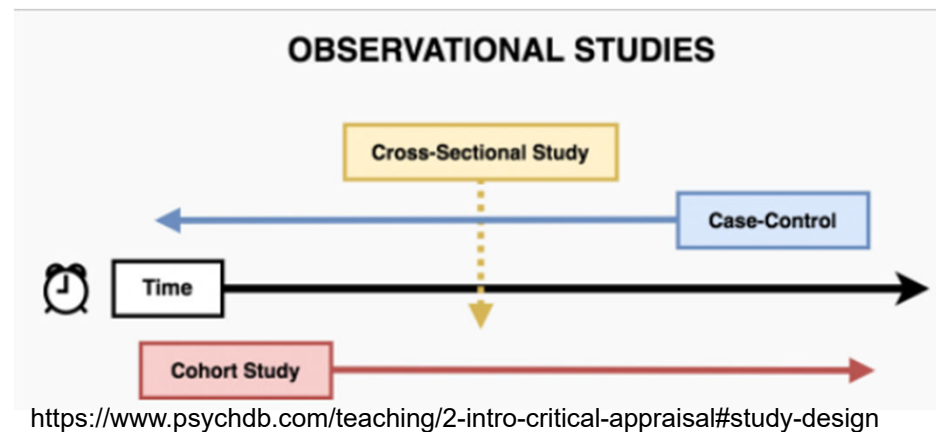
Explanatory Variable:
Amount of protein → Outcome variable:
supplement | Weight at 3
months

- Causality has a direction, i.e. the explanatory variable causes some effect in the outcome, not the other way around

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

1. Observational vs. Designed Experiments

- All primary research is either an **observational** [**study**] or a [**designed**] **experiment**

- In **designed experiments,** one or more explanatory variables are manipulated by the experimenter. This allows you to evaluate a causal hypothesis (e.g. "diet affects weight gain").

- In an **observational study,** you cannot conclude anything about causality.

- **Designed experiments** have much stronger causal power than observational studies. There may be strong ethical, or practical reasons stopping you from performing a designed experiment to answer your research question.

- Surveys are a particularly important type of **observational study** which are covered in our Surveys workshops

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Observational studies

– Very important, but not the focus of this workshop

https://www.psychdb.com/teaching/2-intro-critical-appraisal#study-design

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Other lines of evidence to examine causality

Bradford Hill Criteria

1. **Strength** (effect size): A small association does not mean that there is not a causal effect, though the larger the association, the more likely that it is causal.

2. **Consistency** (reproducibility): Consistent findings observed by different persons in different places with different samples strengthens the likelihood of an effect.

3. **Specificity**: Causation is likely if there is a very specific population at a specific site and disease with no other likely explanation. The more specific an association between a factor and an effect is, the bigger the probability of a causal relationship.

4. **Temporality**: The effect has to occur after the cause (and if there is an expected delay between the cause and expected effect, then the effect must occur after that delay).

5. **Biological gradient**: Greater exposure should generally lead to greater incidence of the effect. However, in some cases, the mere presence of the factor can trigger the effect. In other cases, an inverse proportion is observed: greater exposure leads to lower incidence.

6. **Plausibility**: A plausible mechanism between cause and effect is helpful (but Hill noted that knowledge of the mechanism is limited by current knowledge).

7. **Coherence**: Coherence between epidemiological and laboratory findings increases the likelihood of an effect. However, Hill noted that "... lack of such [laboratory] evidence cannot nullify the epidemiological effect on associations".

8. **Experiment**: "Occasionally it is possible to appeal to experimental evidence".

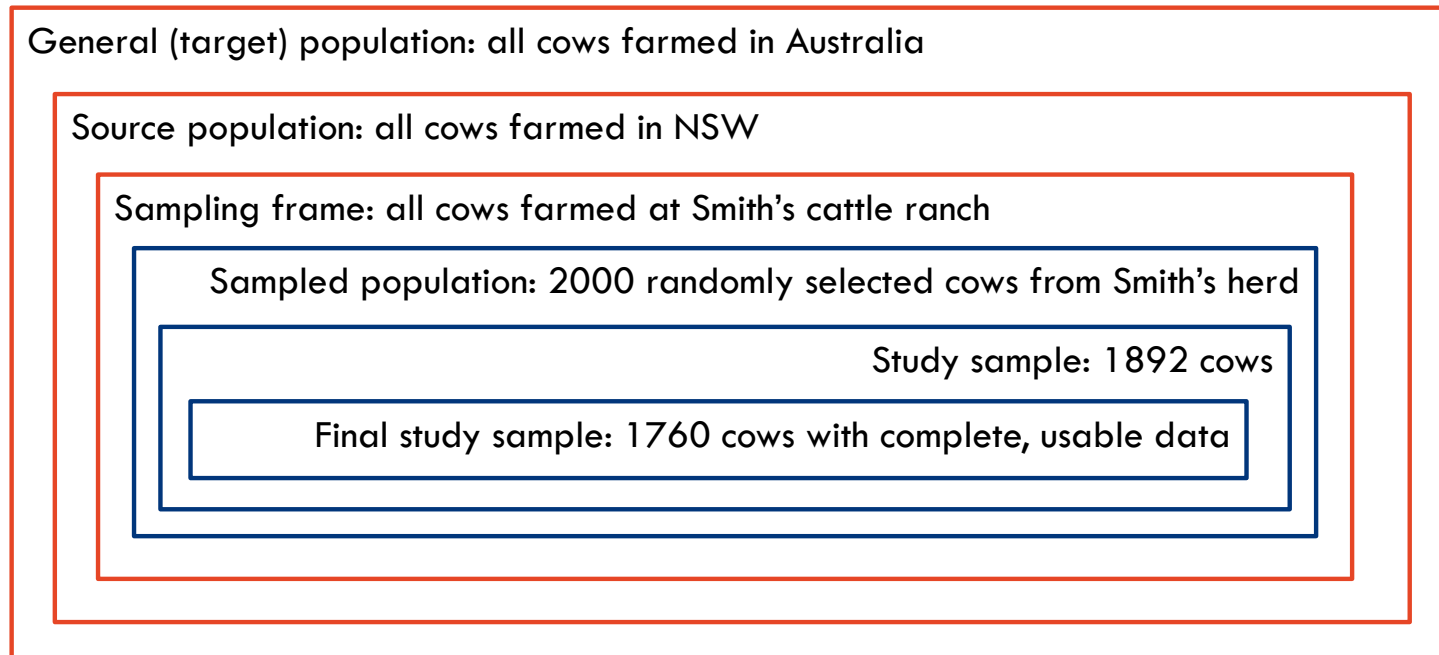9. **Analogy**: The effect of similar factors may be considered.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

1. Identify your target population

– In estimating some population property, we need to consider an important form experimental validity

– External validity is how generalizable the study is to a wider population and depends on the size and representativeness of the sample used

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney
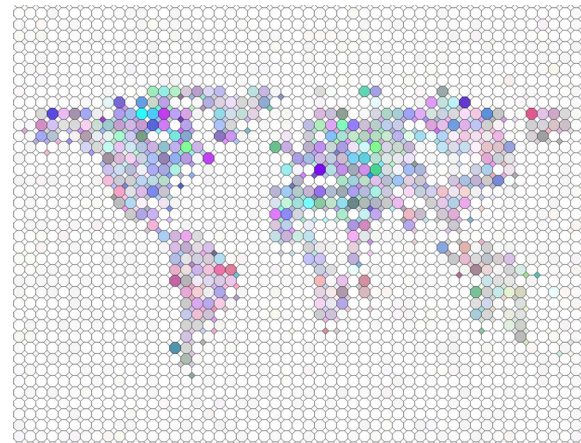
Generalisability to your target population

- When thinking about external validity you need to think across the population and the sample levels. Would the estimates in each of these sets of individuals be the same as the estimates for the larger box?

General (target) population: all cows farmed in Australia

Source population: all cows farmed in NSW

Sampling frame: all cows farmed at Smith's cattle ranch

Sampled population: 2000 randomly selected cows from Smith's herd

Study sample: 1892 cows

Final study sample: 1760 cows with complete, usable data

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

External Experimental validity: Example

Study to evaluate the effect of a feed supplement on the growth of calves

Study Design:

- 2 groups: Std feed and std feed with supplement
- All calves are the same breed - Charolais
- All calves born in the same season
- All male
- All based on Smith's cattle farm

What larger source population does this sample represent?

What general (target) population might we wish to make inferences about?

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

External Experimental validity: Example

Study to evaluate the effect of a feed supplement on the growth of calves

Conclusions will be valid for Charolais bull calves raised under local conditions.

Findings of the study may not be valid for all calves:
- Other breeds
- Female (heifers)
- Grass fed?
- other

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

External Experimental validity: Example

Study to evaluate the effect of a feed supplement on the growth of calves

An expanded study could now include:

- Sex: male, female
- Feed type: Grass and grain
- Breed: Charolais, Hereford
- Climate: temperate, arid

This will expand the external validity of the study to cover a much wider population, but make the study potentially much more difficult to carry out.

Compromise is often necessary. The tension in experimental design is often between what you would *like* to find out and what you feasibly *can* find out in a single experiment.

Generalisability vs. bias

– You may argue that Charolais calves are representative of all calves in a wider population of interest for your outcome of interest

– The danger is when you wish to make inferences within a wider population, without adequately sampling that population or being reasonably sure that the measured individuals in your population are representative

– These issues are where your domain expertise can help you optimise your experimental design

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Generalisability and controlled conditions

- Many experiments are performed under laboratory conditions, perhaps on model organisms or in cells grown *in vitro*

- The question of generalisability always arises when experiments are not performed under 'real world' conditions. Is bias introduced by performing the experiments in a lab compared to the real world?

- On the other hand, controlling conditions often results in far less variability than in the real world (e.g. inbred mouse strains vs. outbred mice). This may be necessary to study subtle effects.

- In practice, a combination of **observational studies** and **designed experiments** are often necessary to fully characterise some particular phenomenon

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

A quick challenge question

– Can you examine causality in the following studies?

- – A randomised control trial where subjects received either a novel drug, or a placebo and their symptoms were examined after 3 weeks

- – An observational (cross sectional) study examining at the rate of disease in subjects of age 60 who routinely took aspirin (for any reason), vs. those that didn't

- – An observational cohort study examining developmental delay in children. Recruited children of the same age and following them over a 10 year period. Comparing outcomes in those who attended pre-school and those that didn't

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

2) Consider analysis of your data in chosen study design

The units in your experiment

So what are the units in your experiment? Understanding the types of units will help you recognise design and analysis considerations

Types of Units (adapted from Lazic)

- Biological/Sampling Unit – is the entity (animal/plant/thing) about which inferences are made

- Experimental Unit – is the entity that is randomly and independently assigned to experimental conditions or treatments

- Observational Unit – is the entity on which measurements are taken

Lazic, Stanley E. *Experimental Design for Laboratory Biologists : Maximising Information and Improving Reproducibility* . Cambridge, United Kingdom: Cambridge University Press, 2016. Print.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

The units in your experiment

Why do these different units matter so much?

When it comes to analysis, you need to deal appropriately with *variability*.

– There will be some amount of variability between your subjects
– Some (usually smaller) amount of variability between measurements of the same subject (if you have them).
– The same issue applies when you have subjects that are similar to each other within a wider sample (e.g. clusters of students in the same class, animals reared on the same farm, mice housed and treated in the same cage etc).

More on this topic later…
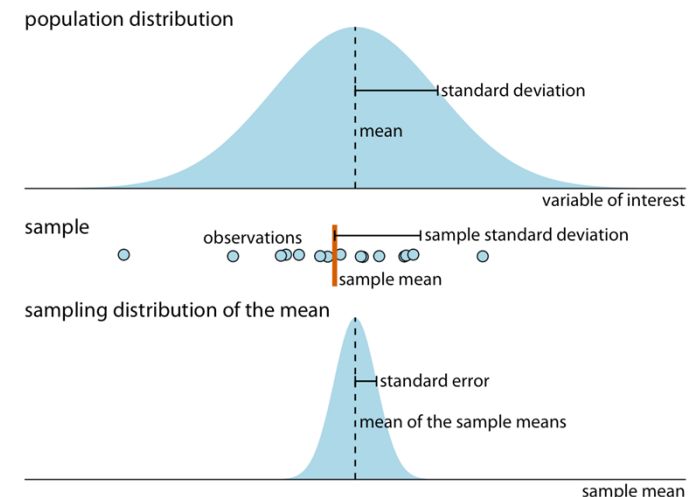
Recall that the variability in the population affects the variability of our estimates – so it matters that different units have different variability. We need methods that take this into account so that we can accurately infer the variability of our estimates (accurate confidence intervals and p-values)

Lazic, Stanley E. *Experimental Design for Laboratory Biologists : Maximising Information and Improving Reproducibility* . Cambridge, United Kingdom: Cambridge University Press, 2016. Print.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Designed Experiments

We will mainly consider study types where you are designing an experiment.

The study conditions are designed and controlled. Often an experimental treatment or intervention is being tested.

To begin our experimental design, we need to have a firm idea of the different variables in our experiment

- Identify and control the explanatory variables:
  - Treatment variables: variables of interest, controlled by the experimenter, e.g. whether subjects receive an intervention (or no intervention)
  - Design variables: other explanatory variables including blocking variables
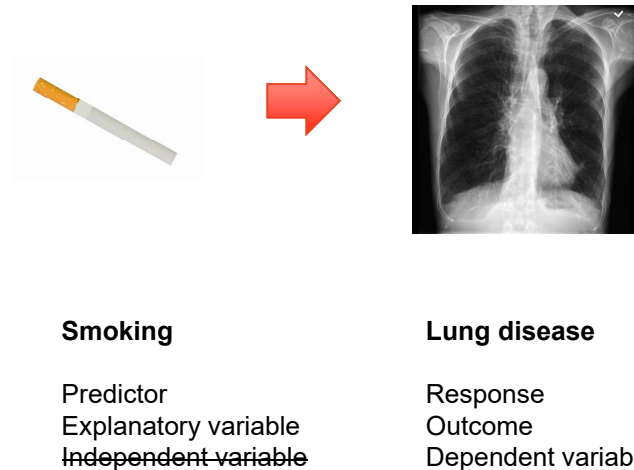- Identify your outcome variable(s)

Designed experiments can identify causal relationships between the explanatory variables and the outcome (response) variables.

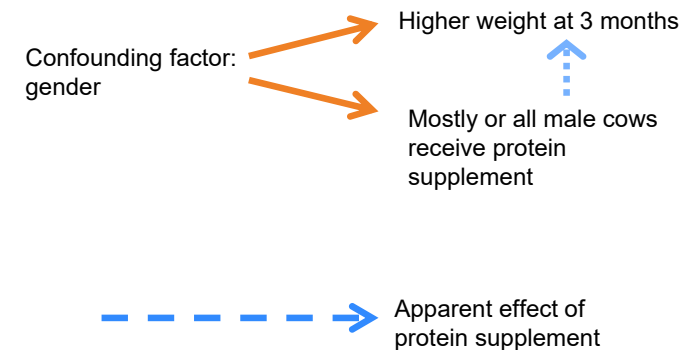# Step 3) Design your experiment

Terminology of experimental design

– We introduced the different types of variables in Research Essentials

– In the language of experimental design, explanatory variables are often referred to as **factors** (e.g. confounding factors)

– The value of the factor for an individual (e.g. smoker, non-smokers) are referred to as levels of the factor

– When we manipulate a factor to test its effect, it becomes a **treatment** variable

From Our Research Essentials Workshop

Smoking

Predictor
Explanatory variable
~~Independent variable~~

Lung disease

Response
Outcome
Dependent variable

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Confounding variables in experimental design

– **Internal validity** is the validity of the experiment for the sample chosen, including proper control of confounding variables

– A *confounding variable* is some variable that may partly or fully explain the apparent (or "uncontrolled effect") of an explanatory variable of interest

– We control for confounders in **designed experiments** by using randomisation and potentially by including other explanatory variables as covariates in our analysis

– In an **observational study,** we don't have experimental intervention, so we instead design our study to sample participants with similar characteristics (e.g. matched pairs), and/or adjust for their differences in analysis

– When confounding is so strong that the effects cannot be disentangled (e.g. *all* male cows receive protein supplement) the study is referred to as **confounded**

Confounding factor: gender → Higher weight at 3 months

→ Mostly or all male cows receive protein supplement

Apparent effect of protein supplement

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Experimental Design – Randomisation

What is randomisation?

Random allocation of treatments to subjects

Why randomise? So we can avoid:

- Systematic bias – e.g. allocating all the drug treatments first, then the placebos
- Selection bias – e.g. subconsciously (or consciously!) choosing healthy patients for the treatment
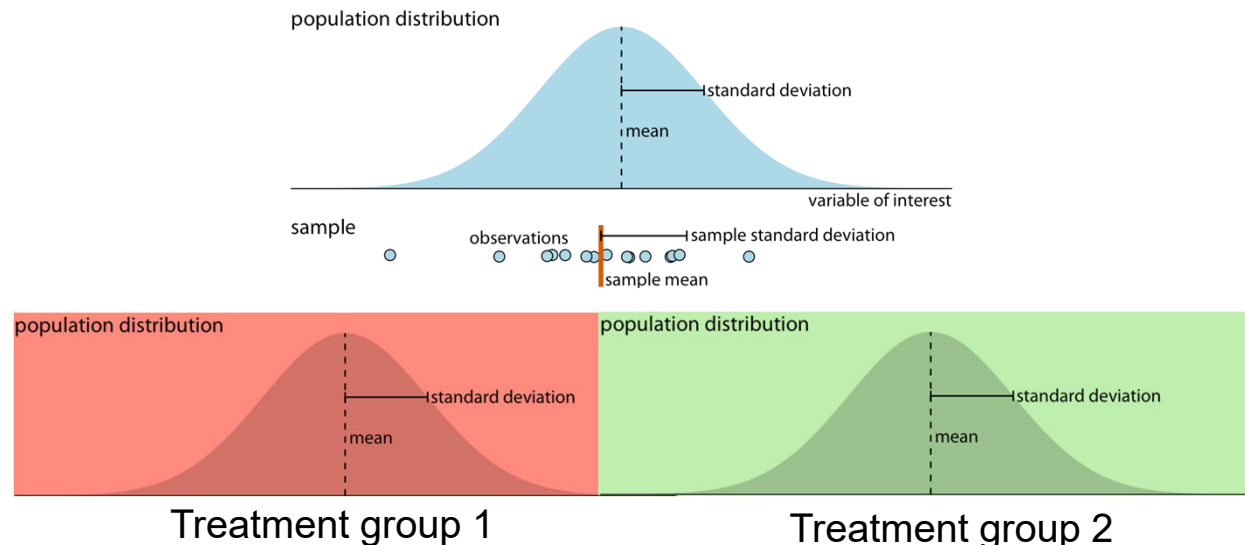- Unknown unknowns – potential confounding factors we don't even know exist.

Randomisation greatly strengthens our causal inference by minimising the effect of confounding variables

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

A philosophical point: randomisation is *not* about perfect balance

Unknown-unknowns cannot be accounted for in an observational study. Some have theorised that all factors including unknowns-unknowns could be perfectly balanced between groups in a randomised experiment. In practice, there are too many unknown-unknowns to achieve perfect balance of all factors across treatment groups with randomisation.

Randomisation does not rely on this perfect balance (which in theory would result in no variability between trials!), instead it allows us to balance the major sources of variability and end up with treatment groups with potentially different means, and individuals that vary to the same extent as individuals in our target population would (after any blocking). Thus we can make as precise estimates as possible in the presence of inevitable randomness

https://errorstatistics.com/2020/04/20/s-senn-randomisation-is-not-about-balance-nor-about-homogeneity-but-about-randomness-guest-post/

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Experimental Design – CRD

Completely Randomised Design (unstructured design)

Example: Evaluate the effect of a feed supplement on the growth of calves

- Suppose that we have no information about the calves (subjects) that we might otherwise use.

- In this case we treat all subjects the same and use randomisation to eliminate allocation biases. (e.g. so the biggest, pushiest cows don't end up in one group)
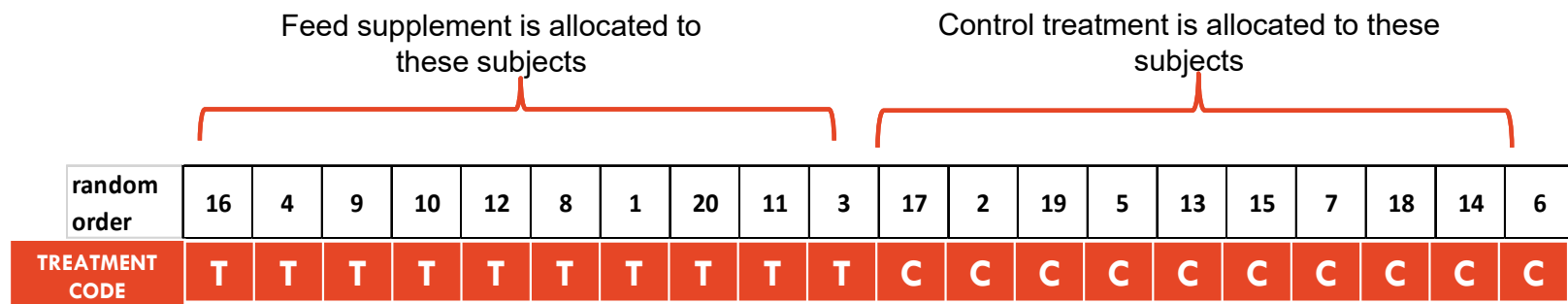
Treatment Group

Randomise!

Control Group

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Completely Randomised Design

A1	B1
16	0.031141412
4	0.041331209
9	0.044095909
10	0.132242434
...	...

- Suppose we have 20 subjects and 2 treatments (T and C)
- Assign an ID number to each subject from 1 to 20

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
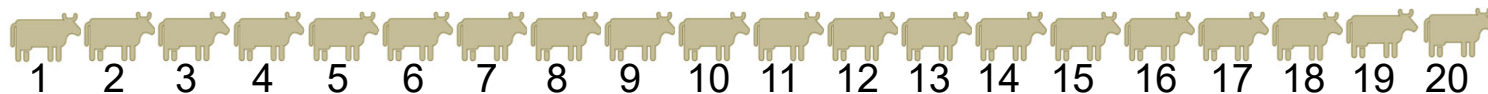
- Generate a randomly ordered sequence of numbers 1 to 20 (eg from Excel)
- In Excel use formulae: with IDs in A1; B1=rand(); copy down 20 rows; sort on B1

Feed supplement is allocated to these subjects

Control treatment is allocated to these subjects

random order	16	4	9	10	12	8	1	20	11	3	17	2	19	5	13	15	7	18	14	6
TREATMENT CODE	T	T	T	T	T	T	T	T	T	T	C	C	C	C	C	C	C	C	C	C

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Completely Randomised Design

How will this experiment be analysed?

We have 2 treatment groups: control and feed supplement

We randomised their treatment ignoring (for the moment) other factors.

Two ways to analyse depending on our design:
1. If we recorded their weight at the beginning of the experiment then we can perform a paired t-test
2. If we didn't record their weight at the beginning of the experiment we can perform an unpaired t-test

For various reasons, option 1. is probably better

Reducing variability between treatment groups

— Recall the two challenges that sampling poses to estimation of the true effect
  — Bias
  — Error

— Randomisation prevents introduction of allocation biases and strengthens causal inference

— What about the known-knowns: factors other than the treatment that are known to affect our outcome, e.g. the gender of the calves? Can we do anything with them?
  — Randomisation also helps with these factors
  — However we can often further minimise their contribution to error by using blocking

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Blocking

- In practice, we may be able to group subjects who are more similar to each other into **blocks**

- A blocking variable is a variable that is thought to affect the outcome, but is typically not of interest to the experimenter. However you may wish to use blocks to increase the representativeness of your sample.

- Balancing your blocking variables within treatment groups will minimise variation in your estimation of the treatment effects

- You may have more than one blocking variable, but it is often only feasible to have a few, so choose those which have the greatest potential effect on your outcome

- Rule of thumb: "Block what you can, randomise what you cannot"

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney
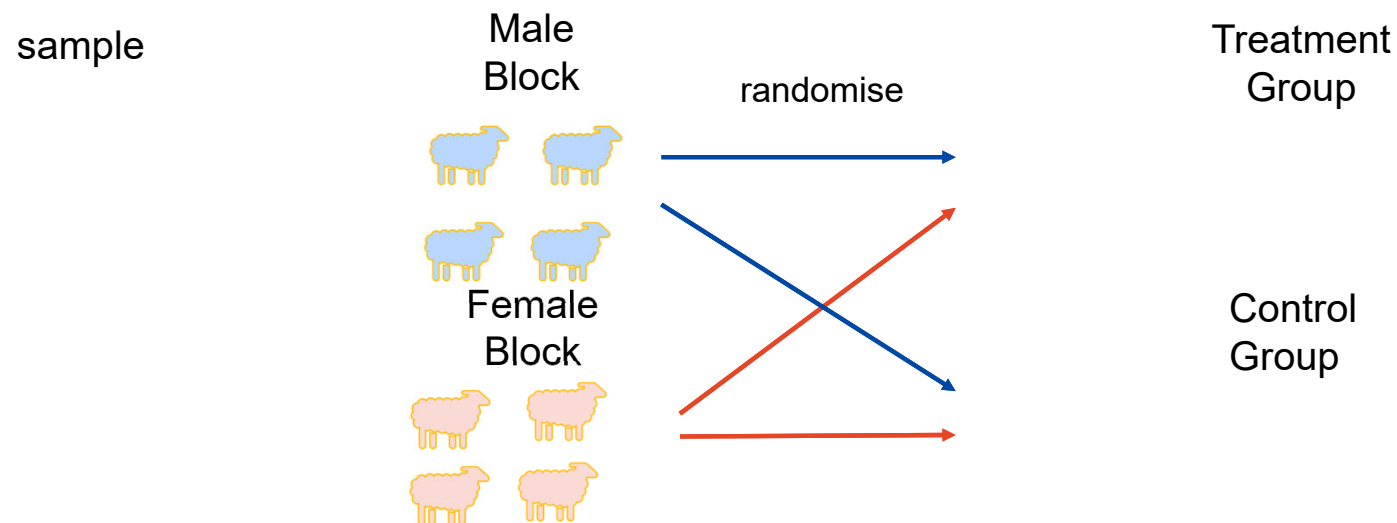
Experimental Design – Randomised Block Design

- **Randomised Block Design (RBD)**
- Example: Evaluate the effect of a feed supplement on the growth of sheep
- Suppose now that we are able to source equal numbers of males and females.
- Use sex as a block variable and randomise within blocks.

sample   Male Block   randomise   Treatment Group

Female Block   Control Group

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Experimental Design – Randomised Block Design

- Randomised Block Design (RBD)
- Example: Evaluate the effect of a feed supplement on the growth of sheep
- Suppose now that we are able to source equal numbers of males and females.
- Use sex as a block variable and randomise within blocks.

sample

Male Block

randomise

Treatment Group

Female Block

Control Group

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney
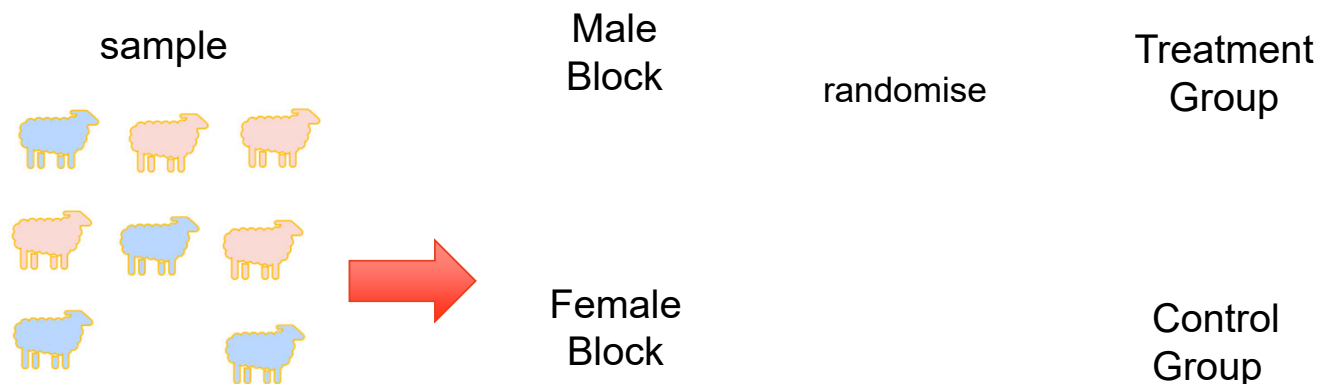
Randomised Block Design

- Sex will be a block variable.
- Male Block

Treatment Code	T	T	T	T	T	C	C	C	C	C
Random order	10	8	2	1	4	9	3	7	6	5

- Female Block

Treatment Code	T	T	T	T	T	C	C	C	C	C
Random order	14	20	19	11	15	18	12	17	16	13

- The allocation is randomised within each block.
  Codes for M 1~10, codes for F 11~20.
- What would be the disadvantage of not blocking for sex in this case?
- How will this experiment be analysed?

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Randomised Block Design

- If we do not think males and females have potentially different responses to treatment then we could use an ANOVA model that adjusts for block and tests for the effect of supplement

  $Y = \beta_o + (supplement)\beta_1 + (male)\beta_2 + \varepsilon$

- If we hypothesise difference in male and female response to protein supplement we use an interaction model

  $Y = \beta_o + (supplement)\beta_1 + (male)\beta_2 + (supplement*male)\beta_3 + \varepsilon$

Main effect of interest: β_1
Blocking variable effect: β_2
Interaction is the additional effect of supplement for male (interaction): β_3

- See linear models series of workshops for more details on analysis using ANOVA

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Latin Square design

Used to create a balanced design with more than one blocking factor.

Example: Growing plants in a greenhouse using different fertilisers

Treatments: Fertiliser A, B, C & D

Row block: shelf position 1,2,3,4

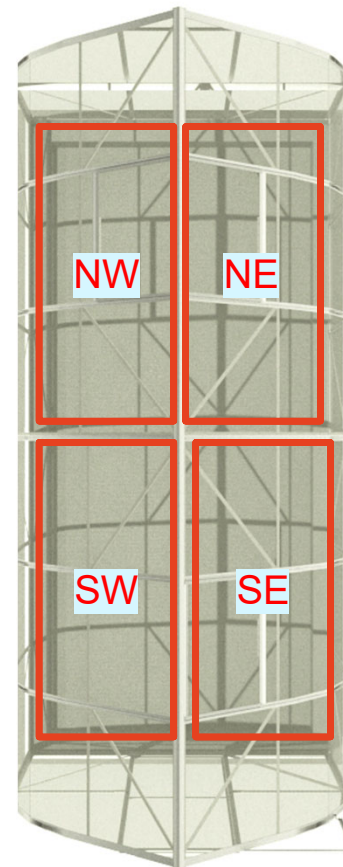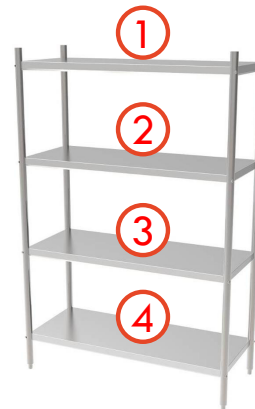Column block: Corner position NE, NW, SE, SW

we can create a **4 x 4 Latin Square design**.

NW   NE

SW   SE
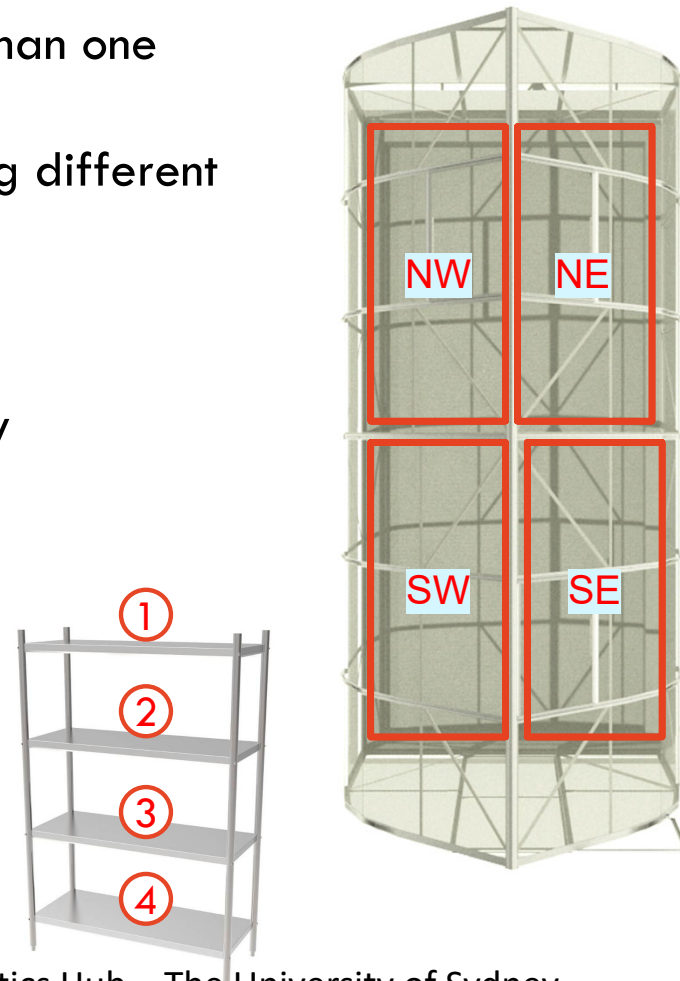
① ② ③ ④

Latin Square design

Example: Growing plants in a greenhouse.

	Column = Shelf position			
	1	**2**	**3**	**4**
NW	A	B	C	D
NE	D	A	B	C
SW	C	D	A	B
SE	B	C	D	A

Row = Greenhouse corner

① ② ③ ④

NW NE

SW SE

Each treatment occurs once per shelf position and once per corner position.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Incomplete block design

– The previous examples worked out well because the size of the block matched the number of treatments (four shelves, four fertilizers)

– The advantage of a complete block design is because all blocks contain all treatments, there is no confounding of treatment effect with the effect of block membership, analysis is simpler and more efficient (less replicates required for same accuracy)

– In reality, our experiments may not work out so nicely. We may want to use blocks that do not have as large a block size as the number of treatments

– In this case we can use a form of incomplete block design, we still aim for balance across the experiment, but cannot achieve it within a single block

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Balanced incomplete block design

- Every *pair* of treatments occurs together in a block the same number of times.

- Example: 4 treatments A, B, C & D

- Block size is only 3 (e.g. 3 shelves, not 4)

Column = Shelf position			
	1	**2**	**3**
Block 1	A	B	C
Block 2	A	B	D
Block 3	A	C	D
Block 4	B	C	D

Don't forget, blocks could be batches, days, cycles, fields, etc

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Completely Randomised Design

Other resources for Randomisation:

- Another nifty random number sequence generator is at: www.random.org/sequences

- Have a look at this video on LinkedIn Learning (through your USyd account, check instructions on Services Portal)

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Factorial designs

Factorial experiments have 2 or more explanatory variables set at 2 or more fixed levels. A common type is the two-level full factorial design

Example: Evaluate the effect of a feed supplement on the growth of calves:

Factor A: Treatment [Supplement/std feed]
Factor B: Feed quantity [low/high]
Factor C: Cattle breed [Charolais/Hereford]

If each *combination* of factor levels forms a single treatment, and the same number of experimental units is allocated to each, this is a 2x2x2 (or 2^3) factorial design.

These designs allow the study of many factors for relatively few runs. Even fewer runs are required with fractional factorial designs. Works best with factors that can easily be controlled.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Questions so far?

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Experimental Design – Blinding

Justice is blind

Why?

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Experimental Design – Blinding

– To avoid bias

– Blind trials (or single blind) – the subject does not know if they are in the treatment or the placebo group

– Double-Blind trials – Both the subject and the technician are not aware of the assigned treatment

– Open trial – All the treatment information is known to the subject and technician/experimenter

– Blinding can reduce or eliminate confounding bias due to conscious or unconscious preferences or expectations.

Experimental Design – Blinding

– Laboratory experiments can also benefit from blinding to prevent bias.

Example 1: Histology cell counts
- Counting cells requires judgement (eg location sampling, recognising cell types)
- The technician should not know the identity of the specimens
- Use an ID code to anonymise the samples. Randomisation of processing order will also help
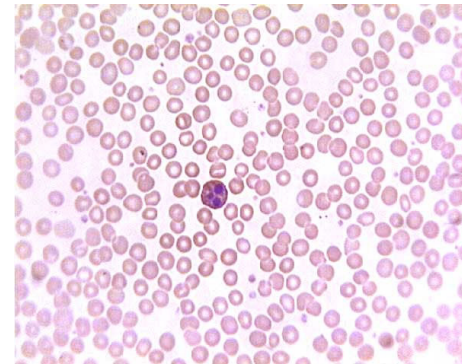
Example 2: Animal behaviour
- Many animals respond to the way they are handled
- The technician should (ideally) not know the identity of the animal's treatment group.

Can you use blinding in your research to guard against unconscious bias?

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney
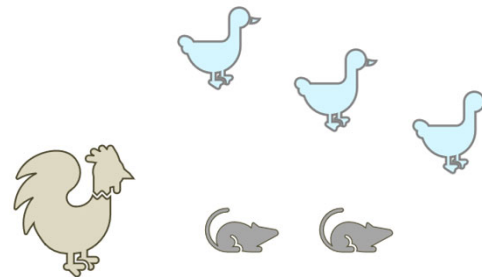
A terminology caution on 'Blocks'

– In Randomised Control Trials you may have heard the term 'block randomisation' to refer to allocation of patients to treatments. This is very different to our Randomised Block Design

– In RCTs we are randomly allocating treatments to participants who enter the study at different times. We use blocks to avoid imbalance in the number in each treatment group during the trial

– E.g. for a two-arm trial with (T)reatment and (C)ontrol we could have six possible blocks of size 4: TCTC, CTCT, TCCT, CTTC, TTCC, CCTT

– If we were recruiting a maximum of sixteen patients, we could randomly choose four of these blocks for our random allocation sequence: TCTC CTCT CTCT CCTT

– To preserve any blinding, you can use the above technique with random block sizes! (Hours of fun)

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Fixed and Random Factors

Fixed Factors

- These are usually the explanatory variables chosen by the experimenter. They have defined levels or categories and we want to quantify the difference between them.

Random Factors

- These are usually incidental to the purpose of the experiment (such as blocking variables).
- The levels of the random factor may be chosen from a larger population of possible values of the variable.
- We don't need to quantify the size of the random effect, we are more interested in using it to partition variance and thereby reduce within group variance.

Discussed in further detail in the Linear Models workshop

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Experimental units

- Types of Units (adapted from Lazic)
- Biological/Sampling Unit – is the entity (animal/plant/thing) about which inferences are made
- Experimental Unit – is the entity that is randomly and independently assigned to experimental conditions.
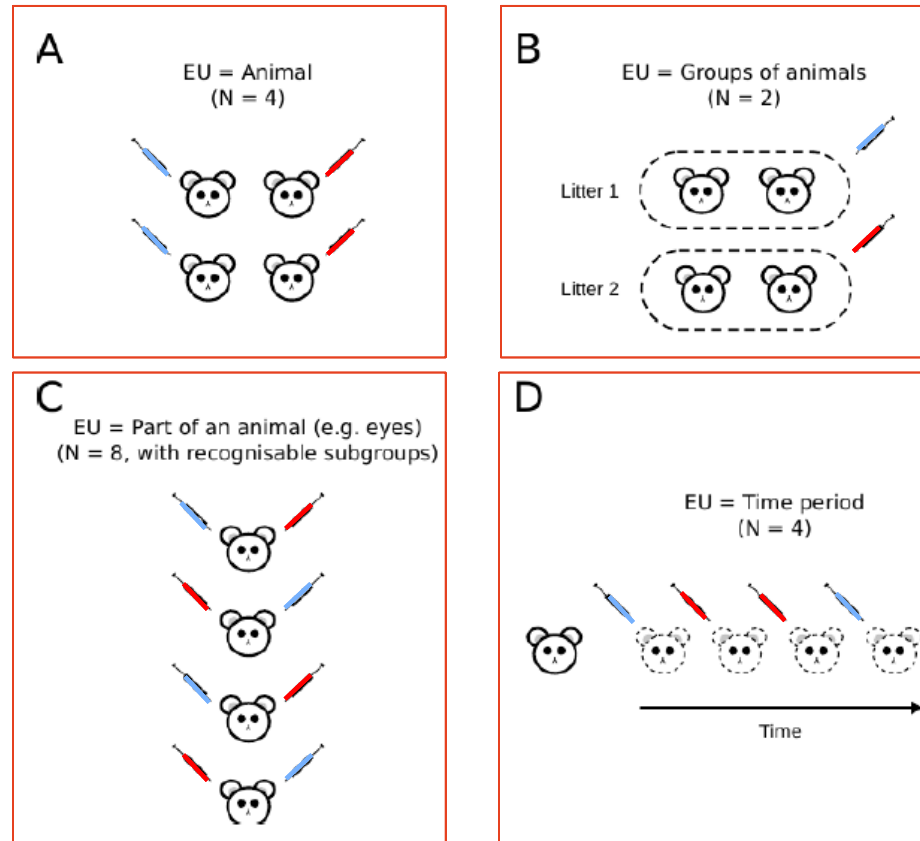- Observational Unit – is the entity on which measurements are taken

Lazic, Stanley E. *Experimental Design for Laboratory Biologists : Maximising Information and Improving Reproducibility* . Cambridge, United Kingdom: Cambridge University Press, 2016. Print.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Experimental Unit leads to N

What about when the treatment is applied to the mother (mare, sow, ewe, etc) and the measurements are carried out on babies in the litter?

Identifying the Biological/Sampling Unit, Experimental Unit and Observational Unit is important (see experimental design workflow)

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Experimental Unit leads to N

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Replication

- True replication occurs when you have multiple independent measurements at the level of each experimental unit.

- Pseudo-replication occurs when repeated measurements on a unit are not independent of each other.

- Technical replication occurs when you take repeated measurements on a unit to increase the precision of that measurement (the measurements are usually averaged)

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Replication

- **Repeated Measures Design (or within-subjects design)**
- Repeated Measures are not technical replicates when they represent another aspect of the same subject/sample, typically observations over time
- Repeated Measures are not independent observations
- There are specific statistical procedures to deal with RM's, discussed in our Linear Models workshops

- It is recommended to calculate the amount of replication required for your experiment using power calculation. The topic of our Power and Sample Size Calculation workshop.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Challenge Question – Sheep vaccine experiment

Research Question: Does the use of a new vaccine result in a different incidence of parasite infection compared to the standard treatment?

I would like 12 sheep in each group (total n = 24)
I have 12 sheep aged 1yr and 12 sheep aged 2yrs

Q: How should you allocate the treatments to the sheep?

<u>a.</u> vaccinate 6 of the younger sheep and 6 of the older sheep with each treatment.

<u>b.</u> vaccinate12 younger sheep with the new vaccine and 12 older sheep with the standard vaccine treatment.

Option b will result in age and treatment being _____

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Challenge Question – Sheep vaccine experiment

Research Question: Does the use of a new vaccine result in a different incidence of parasite infection compared to the standard treatment?

I would like 12 sheep in each group (total n = 24)
I have 12 sheep aged 1yr and 12 sheep aged 2yrs
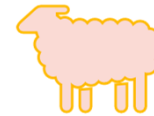
Q: How should you allocate the treatments to the sheep?

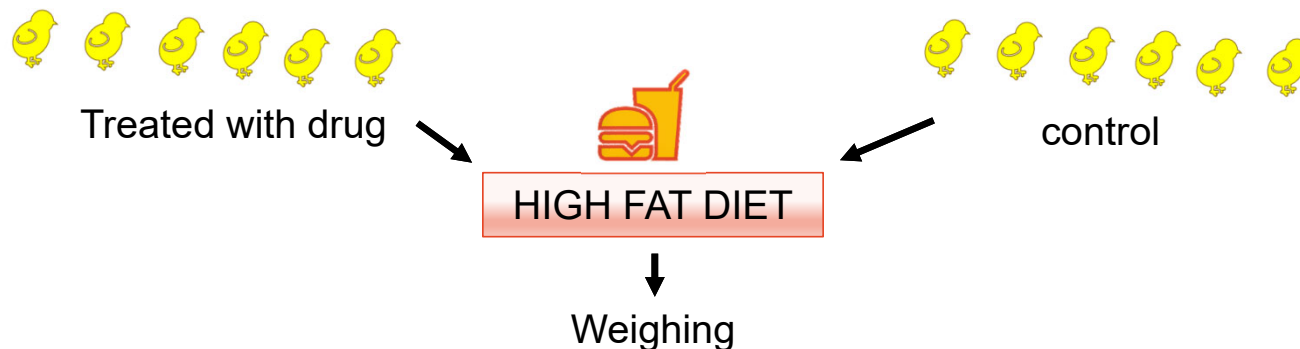a. vaccinate 6 of the younger sheep and 6 of the older sheep with each treatment.

b. vaccinate12 younger sheep with the new vaccine and 12 older sheep with the standard vaccine treatment.

Option b will result in age and treatment being confounded*/aliased

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Scenario 2 – Chicken Drug & Diet experiment

Research Question: Groups of treated and untreated chickens are placed on a high fat diet. What is the effect on weight gain?

- What are the outcome and explanatory variables?
- What are the treatment variables?
- What are the blocking variables?
- What factors should be fixed in the analysis?

Treated with drug

control

HIGH FAT DIET

Weighing

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Scenario 2 – Chicken Drug & Diet experiment

Q. First what are the:

Biological Units	
Experimental Units	
Observational Units	

Treated with drug

HIGH FAT DIET

control

Weighing

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney
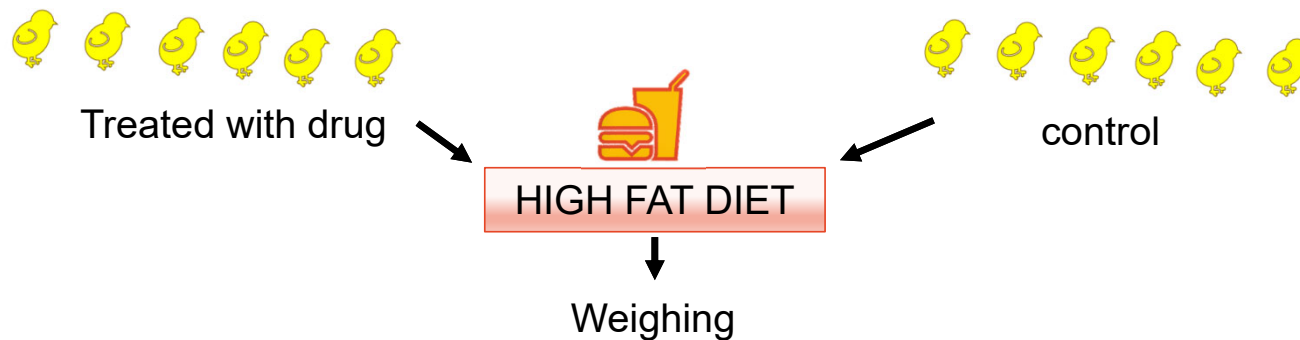
Scenario 2 – Chicken Drug & Diet experiment

Biological Units	chicks
Experimental Units	chicks?
Observational Units	chicks

Treated with drug

control

HIGH FAT DIET

Weighing

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney
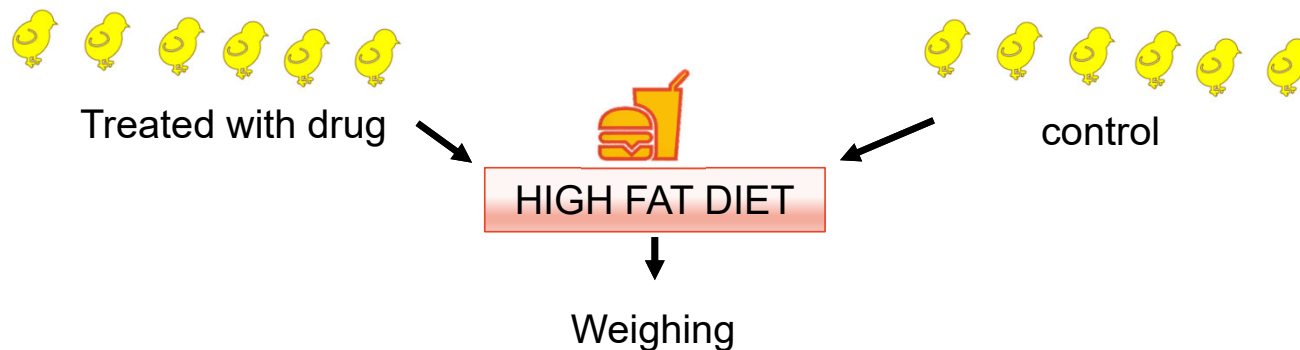
Worksheet - Chicken Drug & Diet experiment

Fill in details

Outcome Variables	Explanatory Variables	Design features	Blocking Factors	Randomisation & blinding

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Worksheet – Chicken Drug & Diet experiment

Outcome variables	Explanatory Variables	Design features	Blocking Factors	Randomisation & blinding
• Weight gain	• Drug treatment (y/n)	• Chick breed • High Fat Diet • Feeding routine • Feeding ad libitum? • Time of day for weighing • Housing – number of chicks per cage?	• Sex? • Chick age, batch, etc	• Drug treatment allocation • Order of handling • Order of weighing

Analytical Design – what statistical test will be used?

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Worksheet – Your research

What does your research experimental design look like?

What are the biological, experimental and observational units?

Biological Units	
Experimental Units	
Observational Units	

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Worksheet – Your research

Outcome Variables	Explanatory Variables	Design features	Blocking Factors	Randomisation & blinding

Experimental Design – what statistical test will be used?

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Experimental design - summary

1. Wherever possible, reduce or eliminate variation due to factors other than the explanatory variable of interest (often treatment variable).
2. Sometimes undesirable variation cannot be avoided due to things beyond your control.
3. Use blocking variables in your design to manage factors that are most likely to cause variation.
4. Use randomisation to prevent bias due to allocation and increase precision in the face of unknown variation outside your control
5. Use replication to improve precision of your estimated effects

- Use the general experimental design workflow in this workshop and note the double arrows between stages of the design: there will often be an iterative process of improvement:
    1. Point out the problems
    2. Discuss the implications
    3. Propose a way forward

- Be alert, not alarmed: your experiment doesn't have to be perfect but it should be well thought out through a process of experimental design

Challenge questions

Q. Types of Units

- _____ Unit – is the entity on which measurements are taken
- _____ Unit – is the entity (animal/plant) about which inferences are made
- _____ Unit – is the entity that is randomly and independently assigned to experimental conditions.

The sample size n = number of _____ units

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Challenge questions

Q. Types of Units

- __observational__ Unit – is the entity on which measurements are taken
- Biological/sampling Unit – is the entity (animal/plant/thing) about which inferences are made
- __experimental__ Unit – is the entity that is randomly and independently assigned to experimental conditions.

The sample size n = number of __experimental__ units

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Further Assistance: Sydney University

SIH
– **1on1 Consults** can be requested on our website (Request Project Help page):
www.sydney.edu.au/research/facilities/sydney-informatics-hub.html OR Google "Sydney Informatics Hub"
– **Training** Sign up to our mailing list to be notified of upcoming training:
mailman.sydney.edu.au/mailman/listinfo/computing_training
  – Research Essentials
  – Experimental Design
  – Power Analysis
  – And more…
– **Hacky Hour**
www.sydney.edu.au/research/facilities/sydney-informatics-hub/workshops-and-training/hacky-hour.html OR Google "Sydney Hacky Hour"

OTHER
– **Open Learning Environment (OLE) courses**
– **Linkedin Learning**: https://linkedin.com/learning/
  – **SPSS** https://www.linkedin.com/learning/machine-learning-ai-foundations-linear-regression/welcome?u=2196204

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Acknowledging SIH

All University of Sydney resources are available to Sydney researchers **free of charge.** The use of the SIH services including the Artemis HPC and associated support and training warrants acknowledgement in any publications, conference proceedings or posters describing work facilitated by these services.

The continued acknowledgment of the use of SIH facilities ensures the sustainability of our services.

Suggested wording:

General acknowledgement:

"The authors acknowledge the technical assistance provided by the Sydney Informatics Hub, a Core Research Facility of the University of Sydney."

Acknowledging specific staff:

"The authors acknowledge the technical assistance of (name of staff) of the Sydney Informatics Hub, a Core Research Facility of the University of Sydney."

For further information about acknowledging the Sydney Informatics Hub, please contact us at **sih.info@sydney.edu.au**.

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

We value your feedback

–   I will email you the the link to an anonymous survey. Prefilled with today's date and the session you attended.

–   It only takes a few minutes to complete (*really!*)

–   Completing this survey is another way to help us keep providing these workshop resources free of charge

Unhappy   Happy

1   2   3   4   5   6   7   8   9   10

Happiness rating

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney

Other resources

Books on Experimental Design

- "The Design of Experiments" by Fisher, Ronald Aylmer, 1935.
- "Experimental Design for Laboratory Biologists: Maximising Information and Improving Reproducibility" by S.E. Lazic
- "Statistics for Experimenters" by Box, Hunter & Hunter
- Interactive E-book on Experimental Design http://cast.massey.ac.nz

Books on Causality

- "The Book of Why" by Judea Pearl (interesting ideas on causality, confounding, approaches to data)

Books on Bias and Statistical thinking

- "Thinking, Fast and Slow" by Daniel Kahneman

Statistical Consulting Unit – Sydney Informatics Hub – The University of Sydney