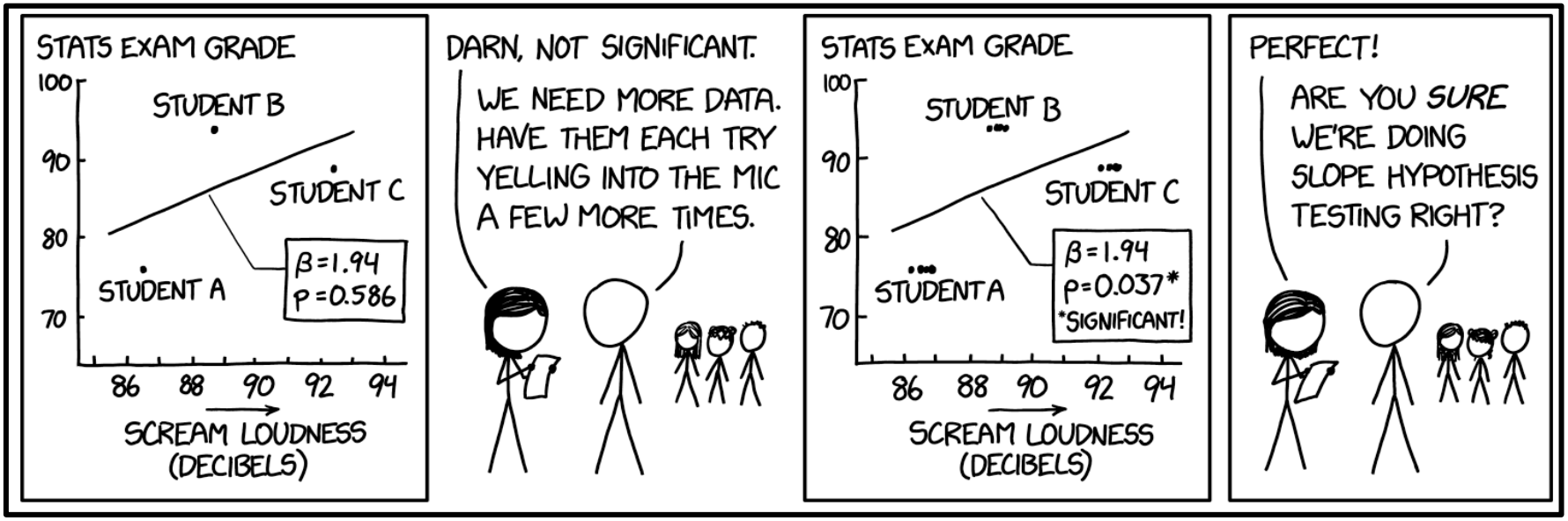
Peer-Feedback Survey
1 survey every Tuesday (constant link)

https://ubc.ca1.qualtrics.com/jfe/form/SV_bvLa6xMXdRk1j6u

BIOL 501: Experimental Design

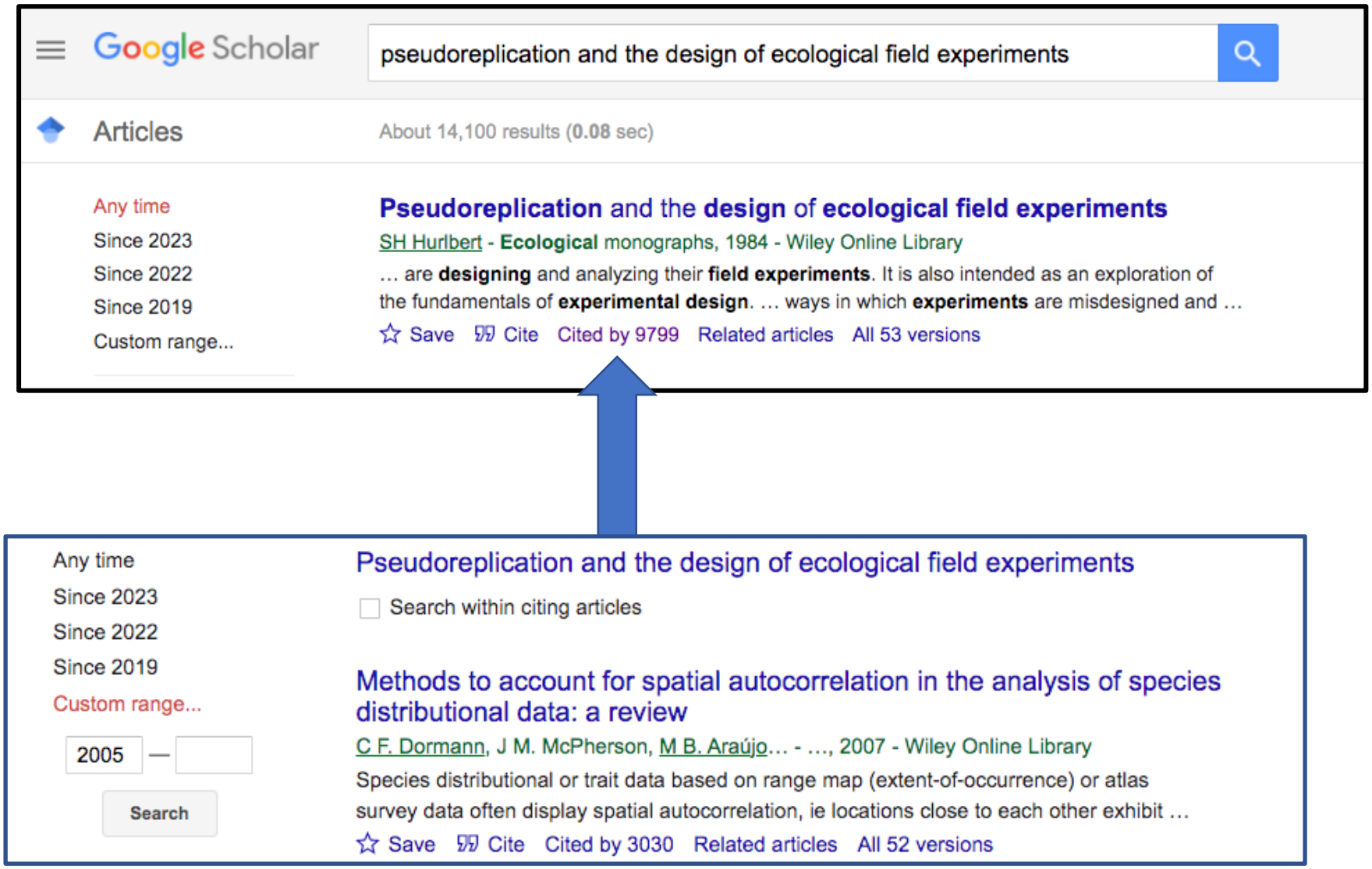


https://xkcd.com/2533/

Get me out, or make a new one

24 Jan. 2023
Dr. Beth Volpov, b.volpov@oceans.ubc.ca; UBC 2022 Term 2

Reverse-look up on Google Scholar



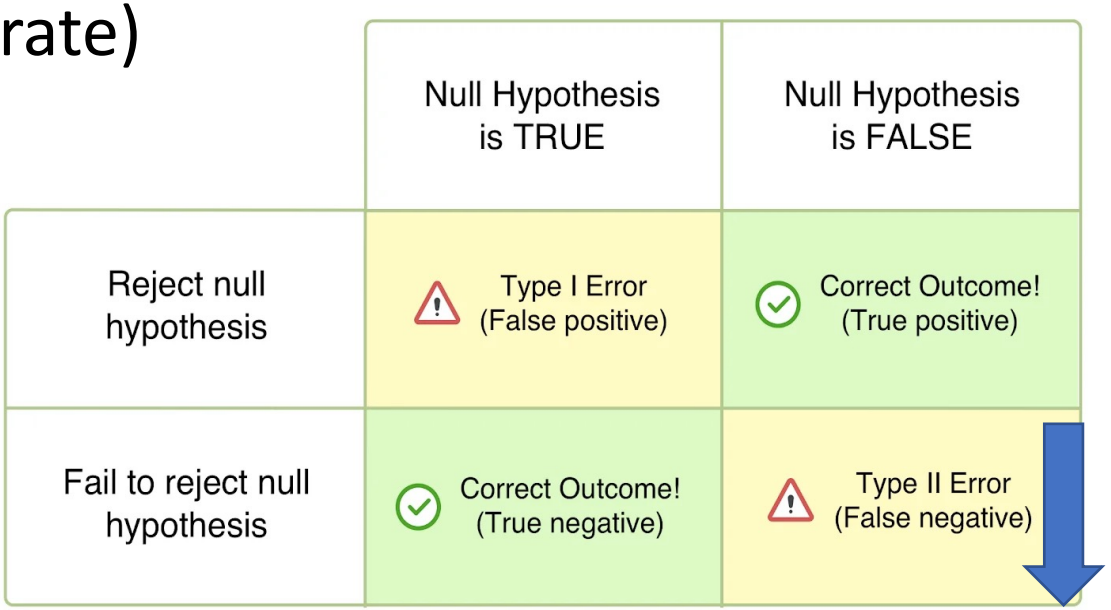
Outline for today

- Sample size and power
- Randomization?
- Minimize bias and effects of sampling error
- Replication and repeated measures
- Balanced vs unbalanced
- Warning of Time as a factor
- Analysis should follow design
- Workshop on Thursday: Plan Experiments
- Optional: for loops practice

Statistical Power and Sample Size

- **Power:** the probability of detecting a true effect, when there actually is one (probability of accepting H_A if it is true)

- High Power→lower probability of making Type II error (low false negative rate)



Problems with Low Power

ANALYSIS

- Overestimates effect size
- Low reproducibility of results
- Wide confidence intervals if < n

Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button[1,2], John P. A. Ioannidis[3], Claire Mokrysz[1], Brian A. Nosek[4], Jonathan Flint[5], Emma S. J. Robinson[6] and Marcus R. Munafò[1]

Abstract | A study with low statistical power has a reduced chance of detecting a true effect, but it is less well appreciated that low power also reduces the likelihood that a statistically significant result reflects a true effect. Here, we show that the average statistical power of studies in the neurosciences is very low. The consequences of this include overestimates of effect size and low reproducibility of results. There are also ethical dimensions to this problem, as unreliable research is inefficient and wasteful. Improving reproducibility in neuroscience is a key priority and requires attention to well-established but often ignored methodological principles.

It has been claimed and demonstrated that many (and possibly most) of the conclusions drawn from biomedical research are probably false[1]. A central cause for this

low sample size of studies, small effects or both) negatively affects the likelihood that a nominally statistically significant finding actually reflects a true effect. We dis-

- Low power studies have a high chance of failing to detect a true effect (false negative rate)

Button et al (2013) *Nature Neuroscienc*

Statistical Power and Sample Size

- Power is often low in biology or when working with animals

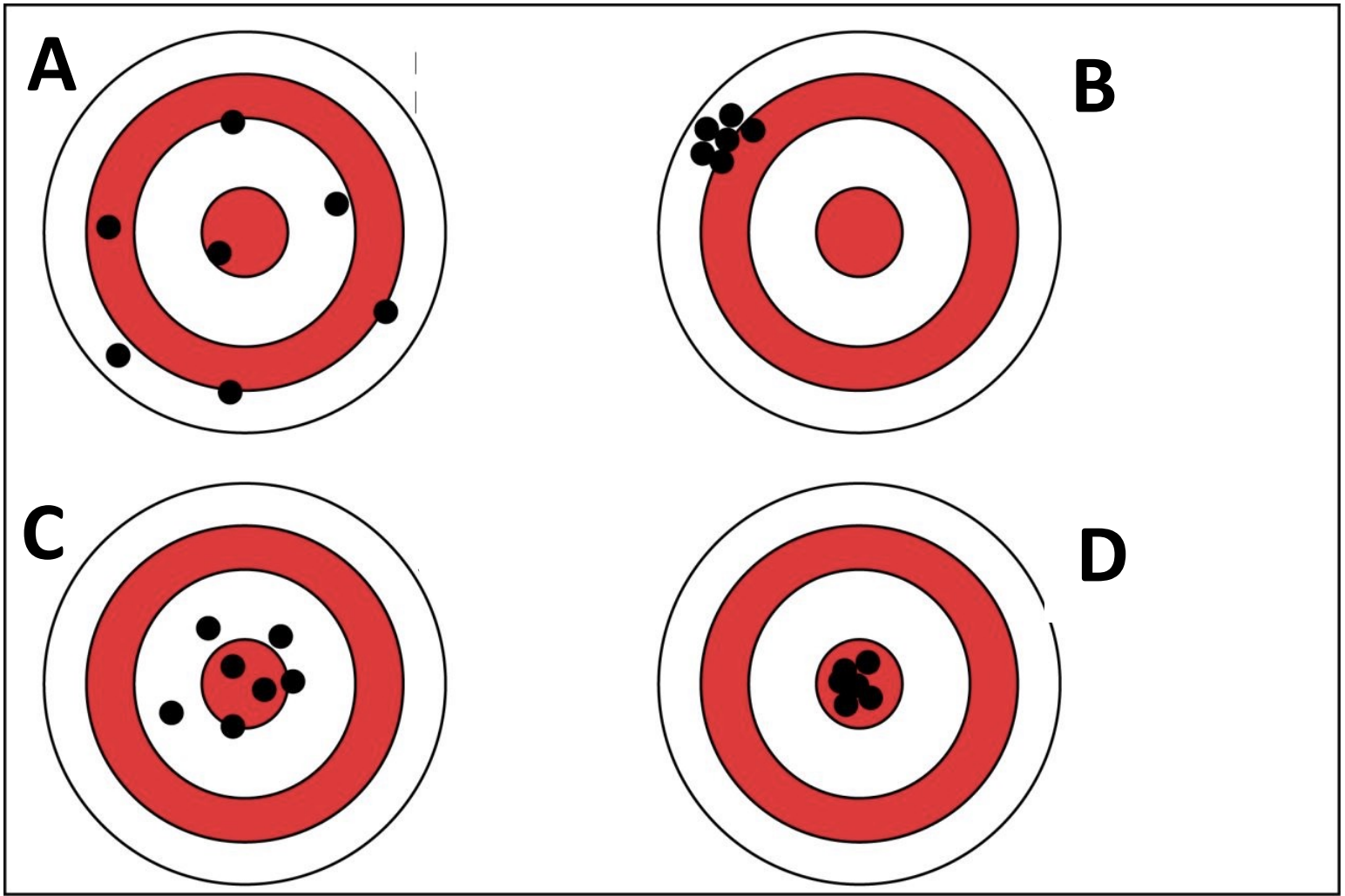
- Plan your sample size (n) wisely
  - Minimizing number of animals used (ethics and cost)

Goals Planning Sample Size (N)

1. Plan for precision
2. Plan for power
3. Compensate for data loss

In Pairs: Review Precision

- For each below, write if the dart board has high or low accuracy and precision



Sample size: plan for precision

- Choose a sample size big enough to yield narrow confidence intervals when estimating an effect
- Narrow CI indicates effect is estimated with high precision

Sample size: plan for power

- Choose a sample size high enough to yield high probability of rejecting null hypothesis (≥ 80%)

Power

- Use R to simulate data before you collect it to help plan sample size (workshop this week)

- Need to estimate key quantities to do power analysis
- Make an educated guess on parameters from literature
  - **But** may be biased if from low power studies→expect real effects to be smaller

- Consider a pilot study before the larger experiment to get information for a power analysis

Power Analysis Packages

- Pwr() package for standard experimental designs
- Use R to simulate data before you collect it to help plan sample size (workshop Thurs)

```
#calculate power of detecting sig. slope LME model simulated data
library(nlmeU)
Pwr(model1,L=c(”x.variable"=1),alpha = 0.05)
```

Discussion: Compensate for data loss

Did you loose data during your thesis? Do you have a plan to compensate for data loss?

Discussion: Your thesis sample size

Why not collect a huge sample size?

What restrictions do (did) you have on your sample size for your thesis data?

What if your sample size is limited?

- Design your study to maximize available sample size
- Do a power analysis and/or pilot study
- Use repeated measures among or within (mixed-effects models)

Goals Planning Sample Size (N) Summary

1. **Plan for precision.** Choose a sample size big enough to yield narrow confidence intervals when estimating an effect
2. **Plan for power.** Choose a sample size high enough to yield high probability of rejecting null hypothesis (≥ 80%)

3. **Compensate for data loss.** Starting sample sizes should account for data loss.

Randomization

- Treatments assigned to units at random, such as by flipping a coin or using random numbers.
- “Haphazard” assignment has repeatedly been shown to be non-random and prone to bias.

Randomization

- Tease apart effects of explanatory variables from those of confounding variables
- Randomization doesn't eliminate the variation contributed by confounding variables. It eliminates only their correlation with treatment.
  - i.e. confounding variables will only be associated with treatments by chance
- Randomization **breaks the association between possible confounding variables and the explanatory variable**, allowing the causal relationship between the explanatory and response variables to be assessed.

Reducing Bias and Sampling Error: lessons learned from clinical trials

OPEN ACCESS Freely available online

PLOS | BIOLOGY

Perspective

Whole Animal Experiments Should Be More Like Human Randomized Controlled Trials

Beverly S. Muhlhausler[1]*, Frank H. Bloomfield[2,3,4], Matthew W. Gillman[5,6]

1 FOODplus Research Centre, School of Agriculture Food and Wine, The University of Adelaide, Australia, **2** Liggins Institute, University of Auckland, Auckland, New Zealand, **3** Department of Paediatrics: Child and Youth Health, University of Auckland, Auckland, New Zealand, **4** Gravida, National Centre for Growth and Development, New Zealand, **5** Obesity Prevention Program, Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts, United States of America, **6** Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, United States of America

The design of clinical trials has been refined because the cost of making a mistake with human subjects is so high.

PLOS Biology, 1 February 2013 | Volume 11 | Issue 2 | e1001481

Reducing Bias and Sampling Error: lessons learned from clinical trials

- To reduce *bias*:
  - Simultaneous control group (participants receiving the placebo).
  - Randomization: treatments were randomly assigned to individuals
  - Blinding: neither the subjects nor the clinicians knew which participants were assigned which treatment

Reducing Bias and Sampling Error: lessons learned from clinical trials

- To reduce the *effects of sampling error*, the experiment included:
  - Replication: multiple independent subjects.
  - Blocking: subjects were grouped according to country, yielding a repetition of the same experiment in different settings (“blocks”).
  - Balance: the number of participants was nearly equal in the two groups within every clinic

Importance of a control group

- A study lacking a control group for comparison cannot determine whether the treatment of interest is the cause of any of the observed changes.

- **Lab experiments:** Control subjects should be perturbed in the same way as the other subjects, except for the treatment itself (as far as ethical considerations permit). (e.g. “sham operation/procedures)”,

- **Fieldwork:** applying a treatment of interest may physically disturb the plots receiving it and the surrounding areas, perhaps by trampling the ground by the researchers. Ideally, the same disturbance should be applied to the control plots.

Blinding Applied to Plants and Non-human animals

- Blinding: conceal information from participants/researchers about which subjects receive which treatment.

- Can be incorporated into experiments on nonhuman subjects using coded tags that identify the subject to a “blind” observer without revealing the treatment (who should also measure units from different treatments in random order).

Blinding Review Terminology

- Blinding is the process of concealing information from participants (sometimes including researchers) about which subjects receive which treatment.
- In a *single-blind* experiment, the subjects are unaware of the treatment that they have been assigned. Can be assumed in most non-human studies.
- In a *double-blind* experiment the researchers administering the treatments and measuring the response are also unaware of which subjects are receiving which treatments.
- Blinding prevents subjects and researchers from changing their behavior, consciously or unconsciously, as a result of knowing which treatment they were receiving or administering.

Minimizing effects of sampling error:

- The goal of experiments is to estimate and test treatment effects against the background of variation between individuals (“noise”) caused by other variables.

- One way to reduce noise is to make the experimental conditions constant. Fix the temperature, humidity, and other environmental conditions, for example, and use only subjects that are the same age, sex, genotype, and so on. In field experiments, highly constant experimental conditions might not be feasible.

Replication

- Replication is **not** only the number of plants used, but the number of independent units in an experiment.

- Replicate is the smallest experimental unit to which a treatment is independently applied (Zar 2010)

An “experimental unit” is the independent unit to which treatments are assigned (typically, it is the *unit that is interspersed*).

More details on replication

- An experimental unit might be a single animal or plant if individuals are randomly sampled and assigned treatments independently.
- Or, an experimental unit might be made up of a batch of individual organisms treated as a group, such as a field plot containing multiple individuals, a cage of animals, a household, a Petri dish, or a family.
- Multiple individual organisms belonging to the same unit (e.g., plants in the same plot, bacteria in the same dish, family members) are likely to be more similar to each other, on average, than are individuals in separate units (apart from the effects of treatment).
- Erroneously treating the single organism as the independent replicate when the chamber or field plot is the experimental unit is *pseudoreplication*.
- *Mixed effects models* can be used to analyze such data while avoiding pseudoreplication.

Pseudoreplication

"use of inferential statistics to test for treatments effects with data from experiments where either treatments are not replicated….or replicates are not statistically independent" -Hurlbert 1984

Pseudoreplication stems from the assumption that one has more statistically independent experimental or sampling units than is actually the case.

Pseudoreplication

- Replicate is the smallest experimental unit to which a treatment is independently applied (Zar 2010)
- Psedoreplication is falsy claiming replicates are independent, when they really are not
- *Mixed effects models* can be used to analyze such data while avoiding pseudoreplication.

Problem of analysis, not design. It happens when the structure of the analysis doesn't match that of the experimental design.

Balance

- Balance reduces influence of sampling error on estimation and hypothesis testing
- Mixed-effects models can account for unbalance
- Greater replication is more important than greater balance

Autocorrelation

- **Temporal autocorrelation:** Measurements taken closer together in time are more likely to be related than those taken apart

- **Spatial autocorrelation**: measurments taken closer in space are more likely to be similar than those far apart

- Packages in R can model autocorrelation

Analysis should follow design

- Think about your analysis before you collect data
- Analysis structure should reflect study design
- Avoid pseudoreplication
- Grouped subjects can be incorporated into your analysis using “mixed effects models”

Workshop Thurs: Planning Experiments

- Generating random numbers with rnorm() and sample()
- Use loops in R to simulate data and analyze the data many times
- This will help you generate estimates of power and precision in various study designs
- For loops!
  - Look at “Loop,Repeat” Tab on R Tips page ahead of time
  - Stay after class today if you want to practice for loops

How to find data for Assignment #1

1. Ask labmates, supervisor, other grad students
2. Use online archives
3. Data Thief (stay after class if want demo)
4. Manually enter data
5. Yes it's ok to redo one of your old Figs from undergrad

Make sure there is room to improve on the "Bad" graph. Too little improvement means we can't assign many marks.

Assignment #1: Due 9pm Jan 28th

- Self-assessment with rubric *prior* to turning it in (or try a peer-assessment)
- Annotate heavily so that your script makes sense to us
- Required header on R script

```
#Date last updated:
#Date created:
#Author Name: First Last
#STUDENT NUMBER:
#R version:
#Platform (Mac or PC):
#Project Description: BIOL 501
#Goal of this script: Assignment #1: Improve a graph)
```

- Submit though Canvas as a single pdf with this specific file name "LASTNAME.FIRSTNAME.STUDENT NUMBER.ASSIGNMENT 1.PDF

For Loops
(will need this for workshop)

- [i] is the iteration or counter
- For loop will start at i<-1 by default
- A good way to troubleshoot or create a function is to think of stepping through an individual iteration

R Tips Pages:loop repeat

```
for(i in 1:5){ print("Back off, man, I'm a scientist") }
```

```
#for each iteration [i] starting at 1 and going to 5 do
everything inside the brackets

for(i in 1:5)
{
    print("Back off, man, I'm a scientist")
}
```

R Tips Pages:loop repeat

Specify total iterations by length of vector (I do this one most often focusing on a vector)

```
for(i in 1:length(depth.m))
{
}
```

Specify total iterations by length of vector within dataframe

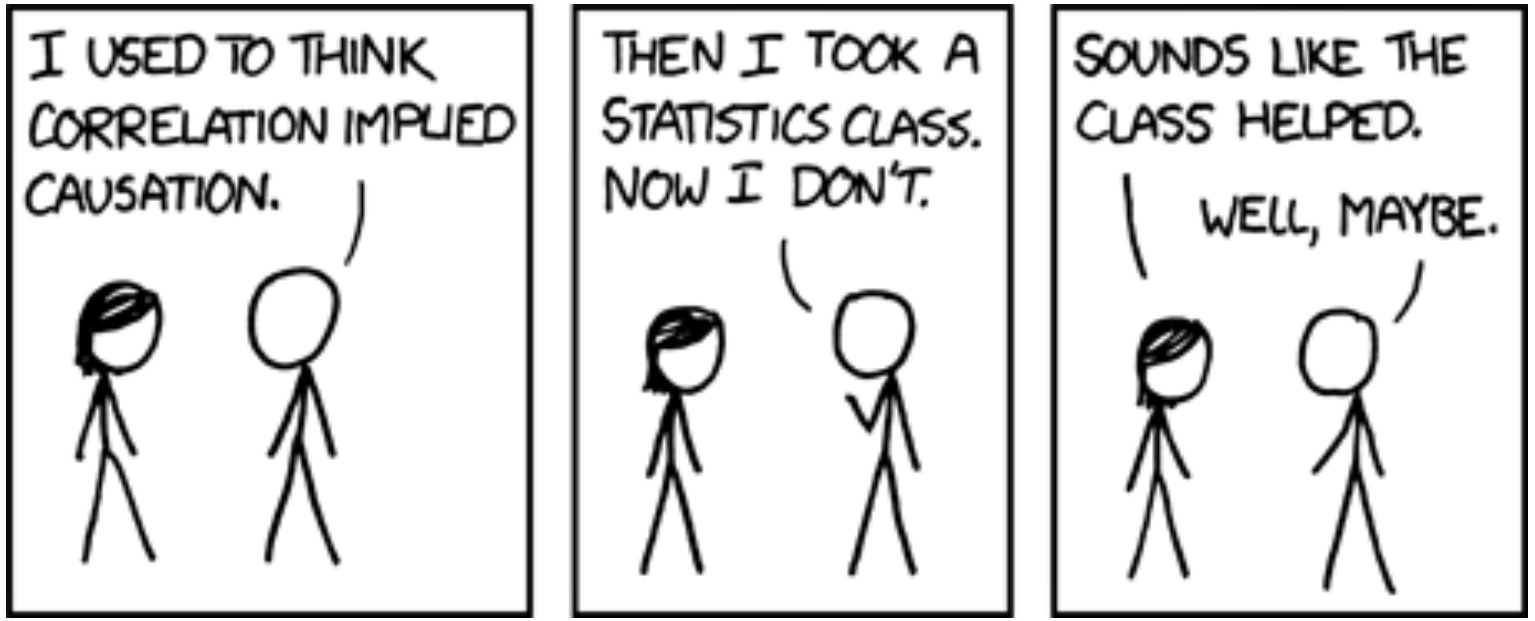
```
for(i in 1:dim(data)[1])
{
}
```

R Tips Pages:loop repeat

- https://www.zoology.ubc.ca/~schluter/R/Loop.html

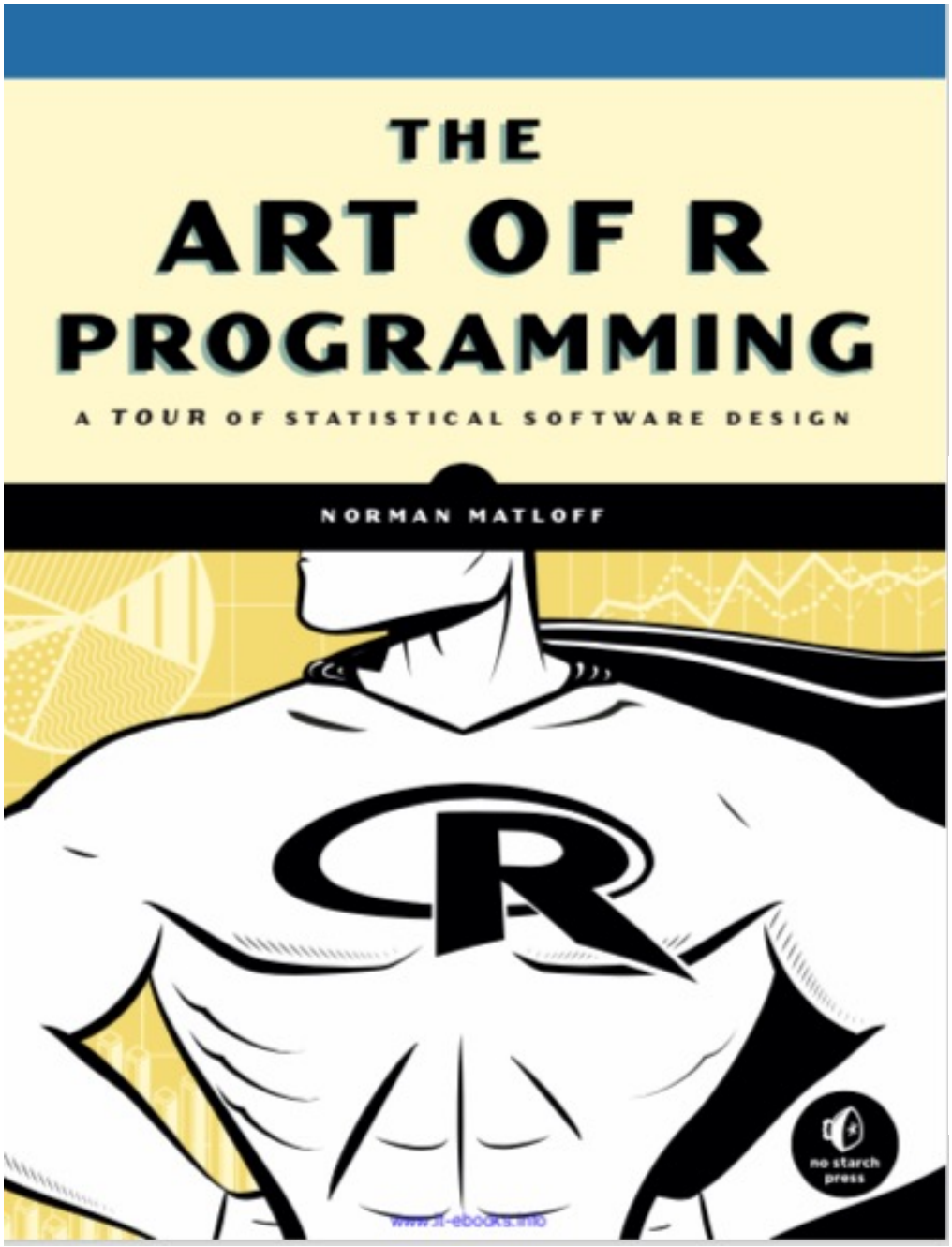
- Stay during office hours today if you want to work through an example on simulated dive depth data for loops

- BIOL 501_Beth's examples of for loops.R



https://xkcd.com/552/

Extra resources



Matloff, Norman. *The art of R programming: A tour of statistical software design*. No Starch Press, 2011.

BRIEF CONTENTS