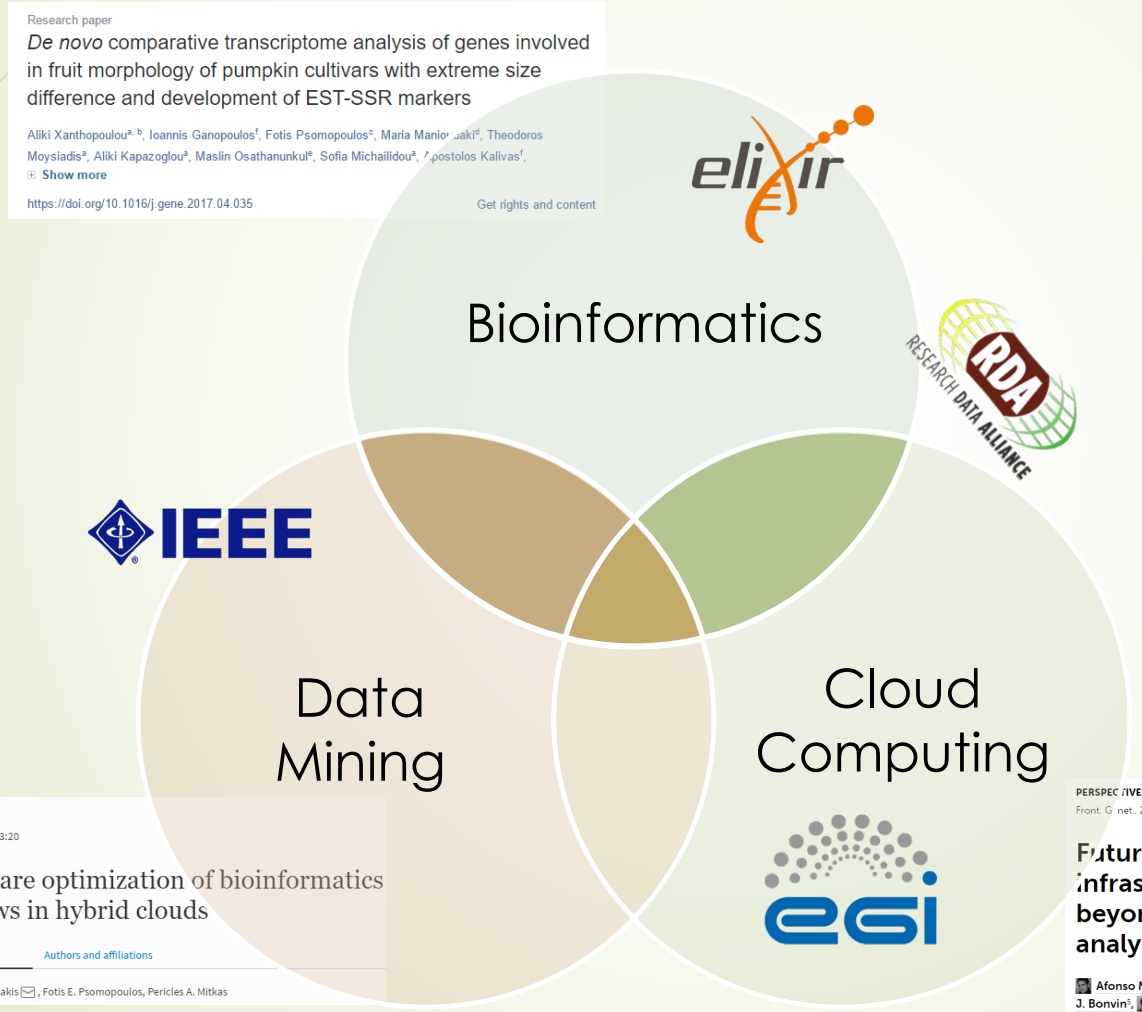
Experiments Design and Analysis

Fotis E. Psomopoulos

CODATA-RDA Advanced Bioinformatics Workshop, 19-23 August 2019, Trieste, Italy

A short intro … to me ☺



- Bioinformatics and Data Mining
  - tools and pipelines to address domain-specific questions
  - genome-aware methods
- Bioinformatics and Cloud Computing
  - workflows and pipelines on cloud infrastructures
  - standardization and reusability

Bioinformatics Group @ INAB | CERTH

Research

- NGS Workflows
- Omics Data Integration
- Data Mining

Training

THE CARPENTRIES

NGS Data Analysis using Cloud Computing (Oct 2015)

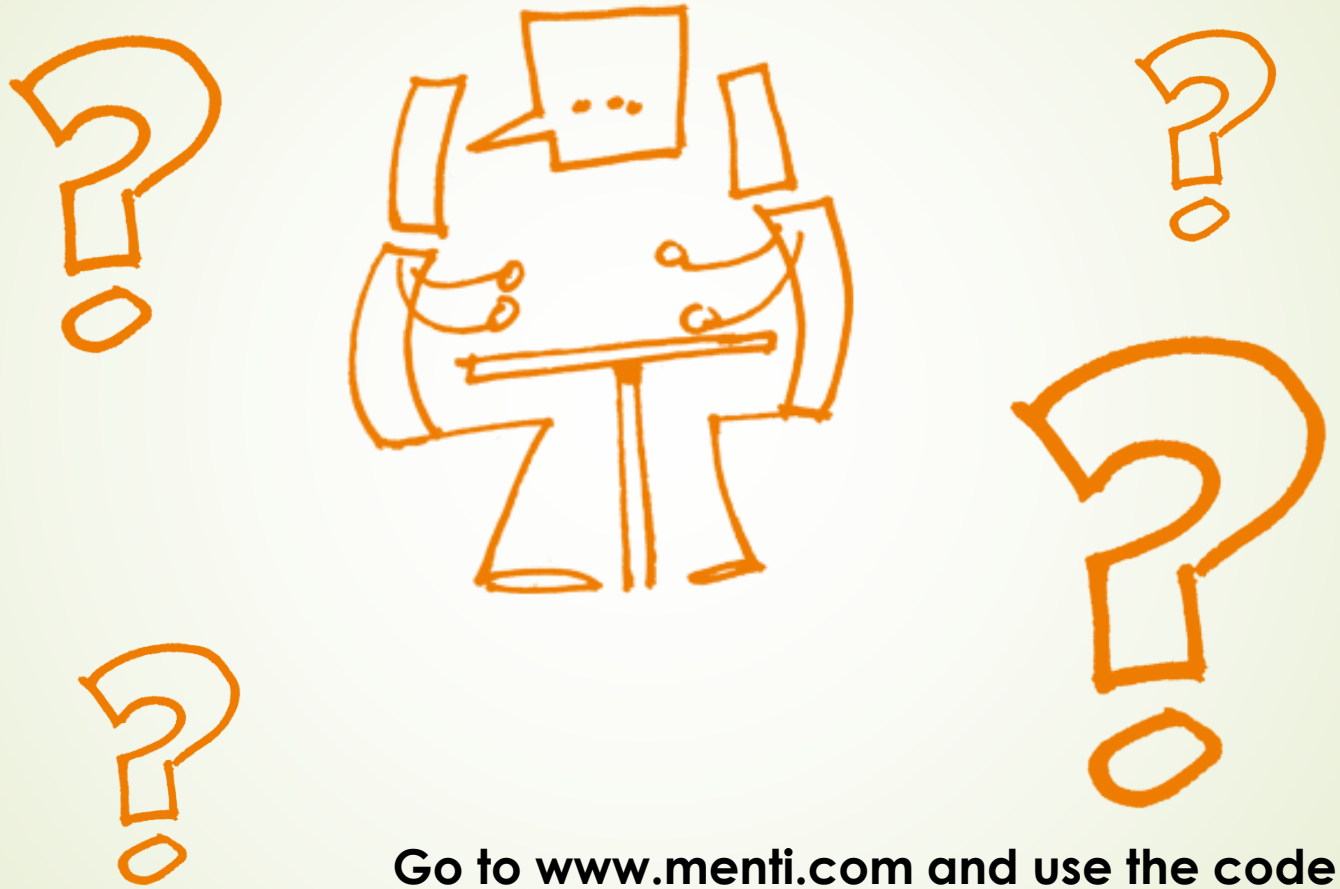
1st Software Carpentry Workshop (Oct 2016)

People

- Maria Kotouza, PhD Student
- Maria Tsayopoulou, PhD Student

CERTH Main Building

Why do we perform experiments?



Go to www.menti.com and use the code 20 11 53

What is an experiment?

An experiment is characterized by the **treatments** and **experimental units** to be used, the way treatments are **assigned** to units, and the **responses** that are measured.

1. Experiments allow us to set up a direct comparison between the treatments of interest.
2. We can design experiments to minimize any bias in the comparison.
3. We can design experiments so that the error in the comparison is small.
4. Most important, we are in control of experiments, and having that control allows us to make stronger inferences about the nature of differences that we see in the experiment. Specifically, we may make inferences about causation.

Components of an Experiment

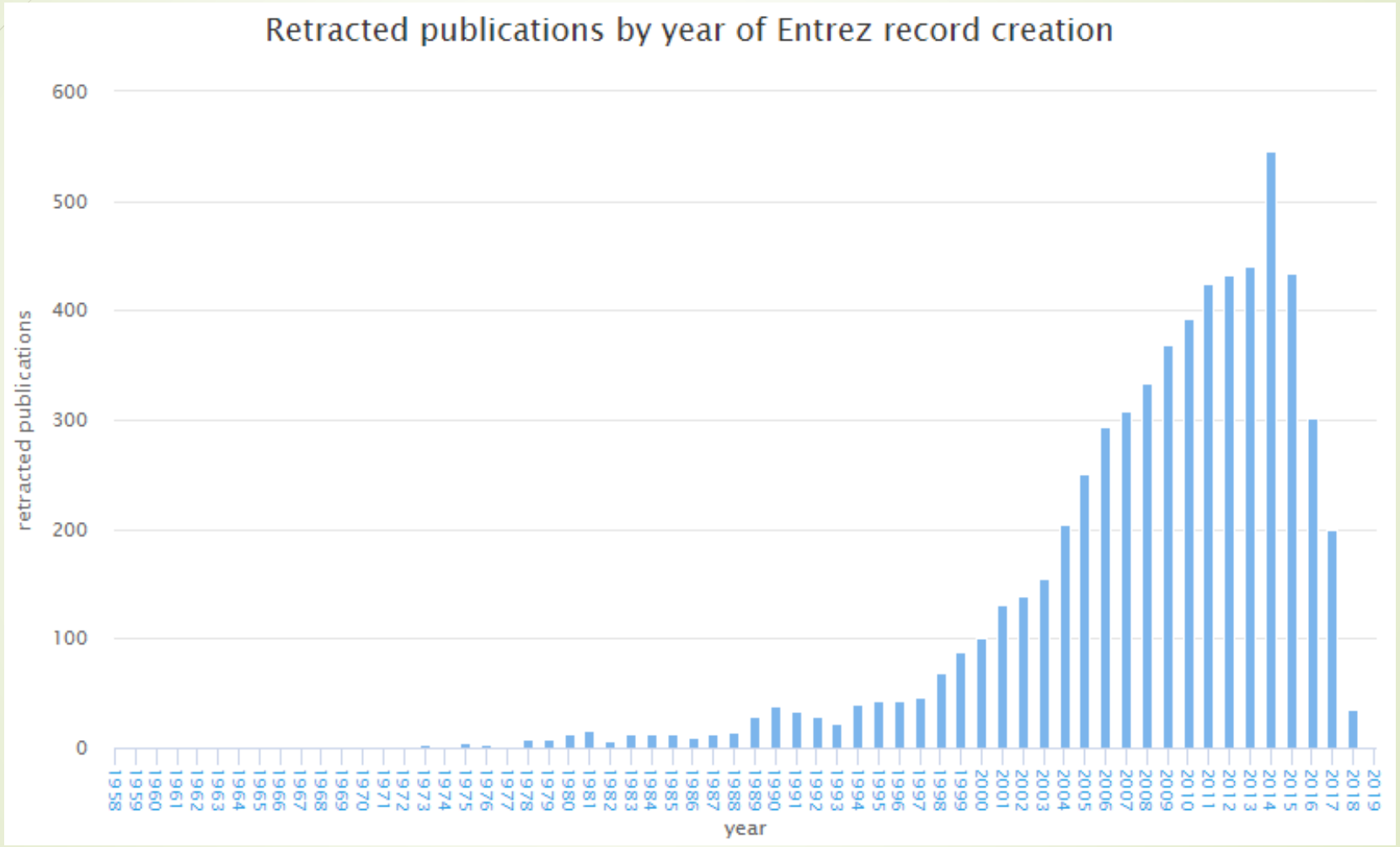
Treatments, units, and assignment method specify the experimental design

- An alternative definition is:
  - "treatment design" is the selection of treatments to be used
  - "experiment design" is the selection of units and assignment of treatments

- Note that there is no mention of a method for analyzing the results.
  - analysis is **not** part of the design
  - However: it is often useful to consider the analysis when planning an experiment.

Why Think About Experimental Design?

Crisis in Reproducible Research



http://neilfws.github.io/PubMed/pmretract/pmretract.html

Consequences of Poor Experimental Design…

- **Cost** of experimentation.
  We have a responsibility to donors!
- **Limited & Precious** material
  esp. clinical samples.
- **Immortalization** of data sets in public databases and methods in the literature.
  Our bad science begets more bad science.
- **Ethical concerns** of experimentation: animals and clinical samples.

Slides adapted from "Designing Functional Genomics Experiments for Successful Analysis", by Rory Stark, 18/09/2017, CRUK-CI

So, what **is** a good experimental design?

Go to www.menti.com and use the code 45 89 48

A good experiment design

- Not all experimental designs are created equal!
- A good experimental design must
  1. Avoid systematic error
  2. Be precise
  3. Allow estimation of error
  4. Have broad validity
- Let's see these aspects one at a time!

Slides adapted from Gary W. Oehlert, "A First Course in Design and Analysis of Experiments", 2010 - ISBN 0-7167-3510-5

1. Design to avoid systematic error

- Comparative experiments estimate differences in response between treatments.

- If an experiment has systematic error, then the comparisons will be biased, no matter how precise our measurements are or how many experimental units we use.

If responses for units receiving **treatment one** are measured with **instrument A** and responses for **treatment two** are measured with **instrument B**,

then we don't know if any observed differences are due to treatment effects or instrument miscalibrations.

2. Design to increase precision

- Even without systematic error, there will be random error in the responses, and this will lead to random error in the treatment comparisons.
- Experiments are precise when this random error in treatment comparisons is small.
- Precision depends on the size of the random errors in the responses, the number of units used, and the experimental design used.

3. Design to estimate error

- Experiments must be designed so that we have an estimate of the size of random error.
- This permits statistical inference:
  - for example, confidence intervals or tests of significance.
- **We cannot do inference without an estimate of error!** Sadly, experiments that cannot estimate error continue to be run.

We will see those in practice later.

4. Design to widen validity

- The conclusions we draw from an experiment are applicable to the experimental units we used in the experiment.
- If the units are actually a statistical sample from some population of units, then the conclusions are also valid for the population.
- Beyond this, we are extrapolating, and the extrapolation might or might not be successful.

We compare two different drugs for treating attention deficit disorder and our subjects are **pre-adolescent boys** from **our clinic**.

- We might have a fair case that our results would hold for pre-adolescent boys elsewhere, but even that might not be true if our clinic's population of subjects is unusual in some way.
- The results are even less compelling for older boys or for girls.

Keeping a common vocabulary

1. Treatments
2. Experimental units
3. Responses
4. Measurement units
5. Randomization
6. Control
7. Factors
8. Confounding
9. Experimental Error
10. Blinding

Terms and concepts (1/5)

1. **Treatments** are the different procedures we want to compare.
   - different kinds or amounts of fertilizer in agronomy
   - different long distance rate structures in marketing
   - different temperatures in a reactor vessel in chemical engineering

2. **Experimental units** are the things to which we apply the treatments.
   - plots of land receiving fertilizer
   - groups of customers receiving different rate structures
   - batches of feedstock processing at different temperatures

Terms and concepts (2/5)

3. **Responses** are outcomes that we observe after applying a treatment to an experimental unit (a measure of what happened in the experiment; we often have more than one response)
   - nitrogen content or biomass of corn plants
   - profit by customer group
   - yield and quality of the product per ton of raw material

4. **Measurement units** (or response units) are the actual objects on which the response is measured. These may differ from the experimental units.
   - (e.g. in different fertilizers on the nitrogen content of corn plants) Different field plots are the experimental units, but the measurement units might be a subset of the corn plants on the field plot, or a sample of leaves, stalks, and roots from the field plot.

Terms and concepts (3/5)

5. **Randomization** is the use of a known, understood probabilistic mechanism for the assignment of treatments to units.
   - Other aspects of an experiment can also be randomized: for example, the order in which units are evaluated for their responses.

6. **Control** has several different uses in design.
   - An experiment is controlled because we as experimenters assign treatments to experimental units. Otherwise, we would have an observational study.
   - A control treatment is a "standard" treatment that is used as a baseline or basis of comparison for the other treatments.
     - This control treatment might be the treatment in common use, or it might be a **null treatment** (no treatment at all).
     - e.g. a study on the efficacy of fertilizer could give some fields no fertilizer at all.

Terms and concepts (4/5)

7. **Factors** combine to form treatments.
   - the baking treatment for a cake involves a given time at a given temperature. The treatment is the combination of time and temperature, but we can vary the time and temperature separately. Thus we speak of a time factor and a temperature factor.
   - Individual settings for each factor are called levels of the factor.

8. **Confounding** occurs when the effect of one factor or treatment cannot be distinguished from that of another factor or treatment.
   - Except in very special circumstances, confounding should be avoided.
   - e.g. planting corn variety A in Minnesota and corn variety B in Iowa. In this experiment, we cannot distinguish location effects from variety effects—the variety factor and the location factor are confounded.

Terms and concepts (5/5)

9. **Experimental Error** is the random variation present in all experimental results.
   - Different experimental units will give different responses to the same treatment, and it is often true that applying the same treatment over and over again to the same unit will result in different responses in different trials.
   - Experimental error does not refer to conducting the wrong experiment or dropping test tubes.

10. **Blinding** occurs when the evaluators of a response do not know which treatment was given to which unit.
   - helps prevent bias in the evaluation, even unconscious bias from well-intentioned evaluators.
   - Double blinding occurs when both the evaluators of the response and the (human subject) experimental units do not know the assignment of treatments to units.

Ok, let's go back to our initial question:

*What **is** a good experimental design?*

A Well-Designed Experiment

- Should have:
  1. Clear Objectives
  2. Focus and Simplicity
  3. Sufficient Power
  4. Randomized Comparisons

- And be:
  1. Precise
  2. Unbiased
  3. Amenable to statistical analysis
  4. Reproducible

Aspects of Experimental Design

- Experimental Factors
- Variability
  1. Sources of Variance
  2. Replicates
- Bias
  1. Confounding factors
  2. Randomization wherever a decision is to be made
     - Controls for **both** measured and unmeasured factors
  3. Controls

Experimental Factors

- Factors: Aspects of Experiment that change and influence the outcome of the experiment
  - e.g. time, weight, drug, gender, ethnicity, country, plate, cage etc.
- Variable type depends on type of measurement
  - Categorical (**nominal**) , e.g. gender
  - Categorical with ordering (**ordinal**), e.g. tumor grade
  - **Discrete**, e.g. shoe size, number of cells
  - **Continuous**, e.g. body weight in kg, height in cm
- Independent or Dependent Variables
  - Independent variable (IV): what **you** change
  - Dependent variable (DV): what changes **due to IV**
  - *"If (independent variable), then (dependent variable)"*

Sources of Variation

- Biological "Noise"
  - Biological processes are inherently stochastic
  - Single cells, cell populations, individuals, organs, species….
  - Timepoints, cell cycle, synchronized vs. unsynchronized
- Technical Noise
  - Reagents, antibodies, temperatures, pollution
  - Platforms, runs, operators
- Consider in advance and control replication required to capture variance

Types of Replication

- Biological Replication
  - In vivo
    - Patients
    - Mice
  - In vitro
    - Different cell lines
    - Re-growing cells (passages)
- Technical Replication
  - Experimental protocol
  - Measurement platform (i.e. sequencer)

How many samples? Why do you need replicates?

- Calculating appropriate sample sizes
  - Power calculations
  - Planning for precision
  - Resource equation
- Power: the **probability** of detecting an **effect** of a specified size if present.
  - Identify and control the sources of variability
    - Biological variability
    - Technical variability
  - Using appropriate numbers of samples (sample size/replicates)
  - Power calculations estimate sample size required to detect an effect *if degree of variability is known*
    - Depends on δ, n, sd, a, H_A
  - If adding samples increases variability, that alone won't add power!

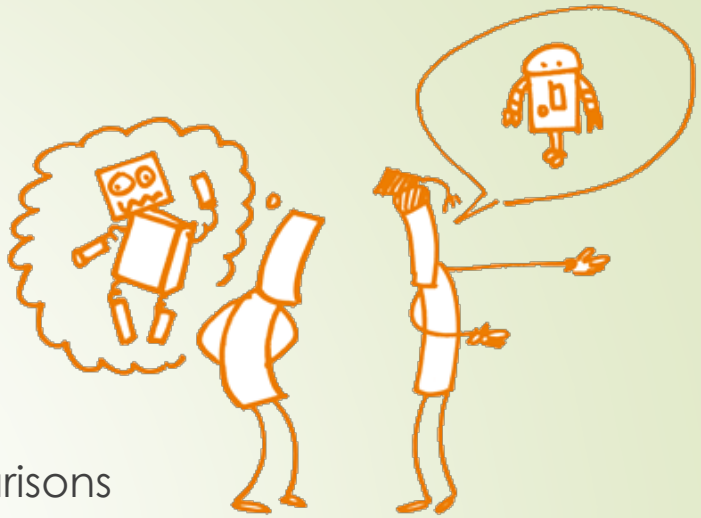
Confounding Factors

- Aka Extraneous, hidden, lurking or masking factors, or the third variable or mediator variable.
- May mask an actual association or falsely demonstrate an apparent association between the independent and dependent variables.
- Hypothetical example would be a study of coffee drinking and lung cancer.

Confounding factors

- Inadequate management and monitoring of confounding factors
  - one of the most common causes of researchers wrongly assuming that a correlation leads to a causality.
- If a study does not consider confounding factors, don't believe it!

Solutions!

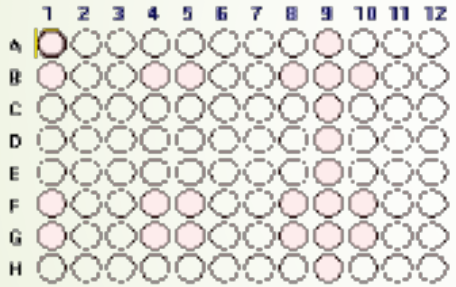


- Randomization
  - Statistical analysis assume randomized comparisons
  - May not see issued caused by non-randomized comparisons
  - Make every decision random not arbitrary
- Blinding
  - Especially important where subjective measurements are taken
  - Every experiment should reach its potential degree of blinding

Technical Confounding Factors: Batch Effects

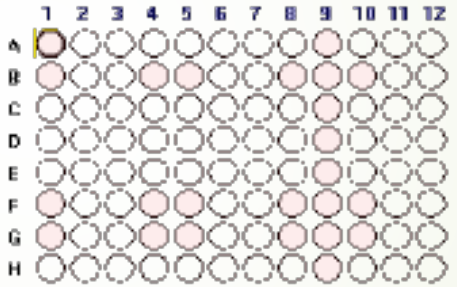
- RNA Extraction

Day 1, Plate 1



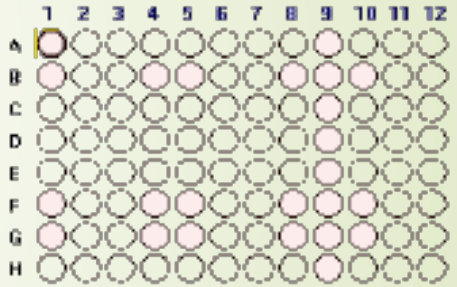
Control

Day 2, Plate 2



Treatment 1

Day 3, Plate 3

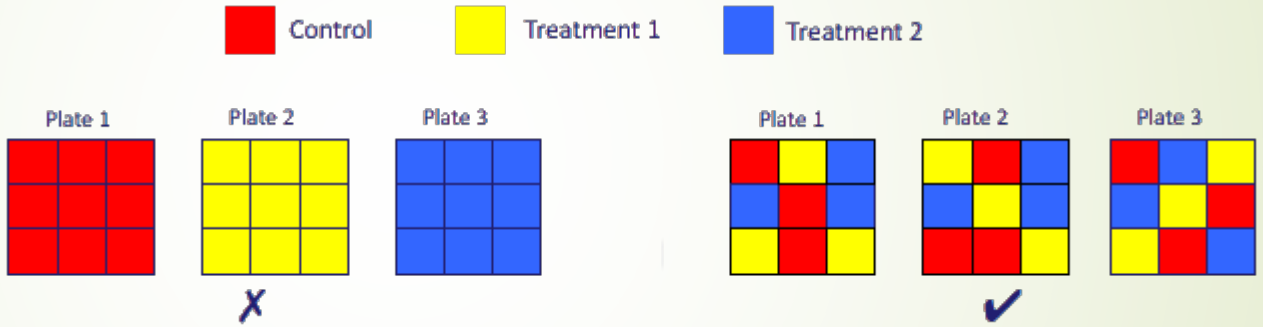


Treatment 2

- The difference between Control, Treatment 1 and Treatment 2 is confounded by day and plate.

Randomized Block Design

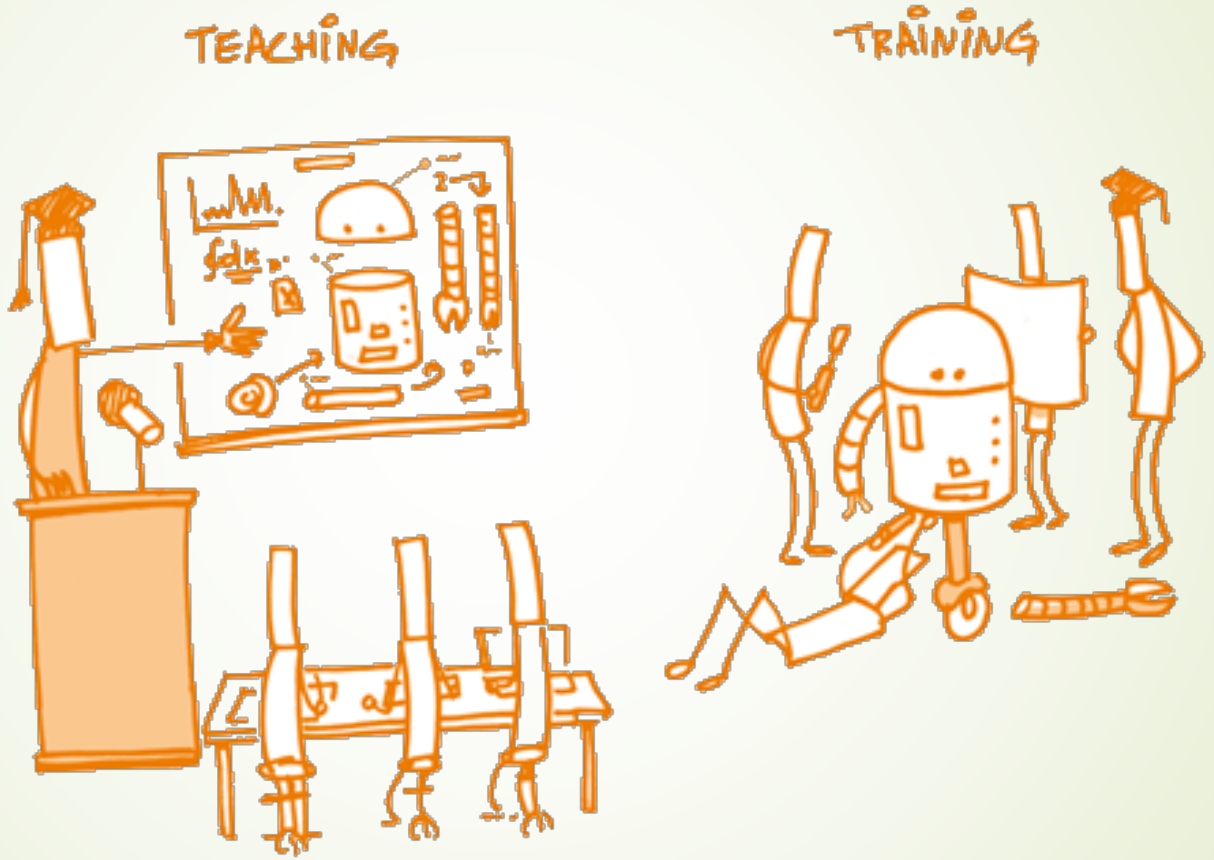
- Blocking is the arranging of experimental units in groups (blocks) that are similar to one another



- RBD across plates so that each plate contains spatially randomized equal proportions of:
  1. Control
  2. Treatment 1
  3. Treatment 2

  controlling plate effects.

All good in theory, but in practice?



Experimental Design Practical Questions I

1. What are your objectives?
2. What are you measuring?
3. What are your primary sample groups of interest?
4. What controls will you use each type of sample group?
5. What constitutes a replicate in this experiment? Are they biological or technical? How many samples/replicates should be collected?
6. Sketch out the design as a matrix, with sample numbers
7. What sample group comparisons (contrasts) will you make with the data? Which gene set(s) will you use for pathway analysis?
8. What are possible confounding factors and sources of bias?

Experimental Design Practical Questions II

9. How will you confirm effective silencing?
10. What information about your experiment should be recorded to help identify any problems should there be any?
11. Will you be multiplexing samples? How will you assign barcodes? Will you use pooled libraries? How many pools? How will samples be assigned to pools?
12. What are the sequencing parameters you need to be aware of (e.g. sequencing type and depth)?
13. What other types of data might be useful to assay, and how might the sequencing parameters need to change to accommodate this?
14. Can you think of any other design related issues that could/should be addressed?

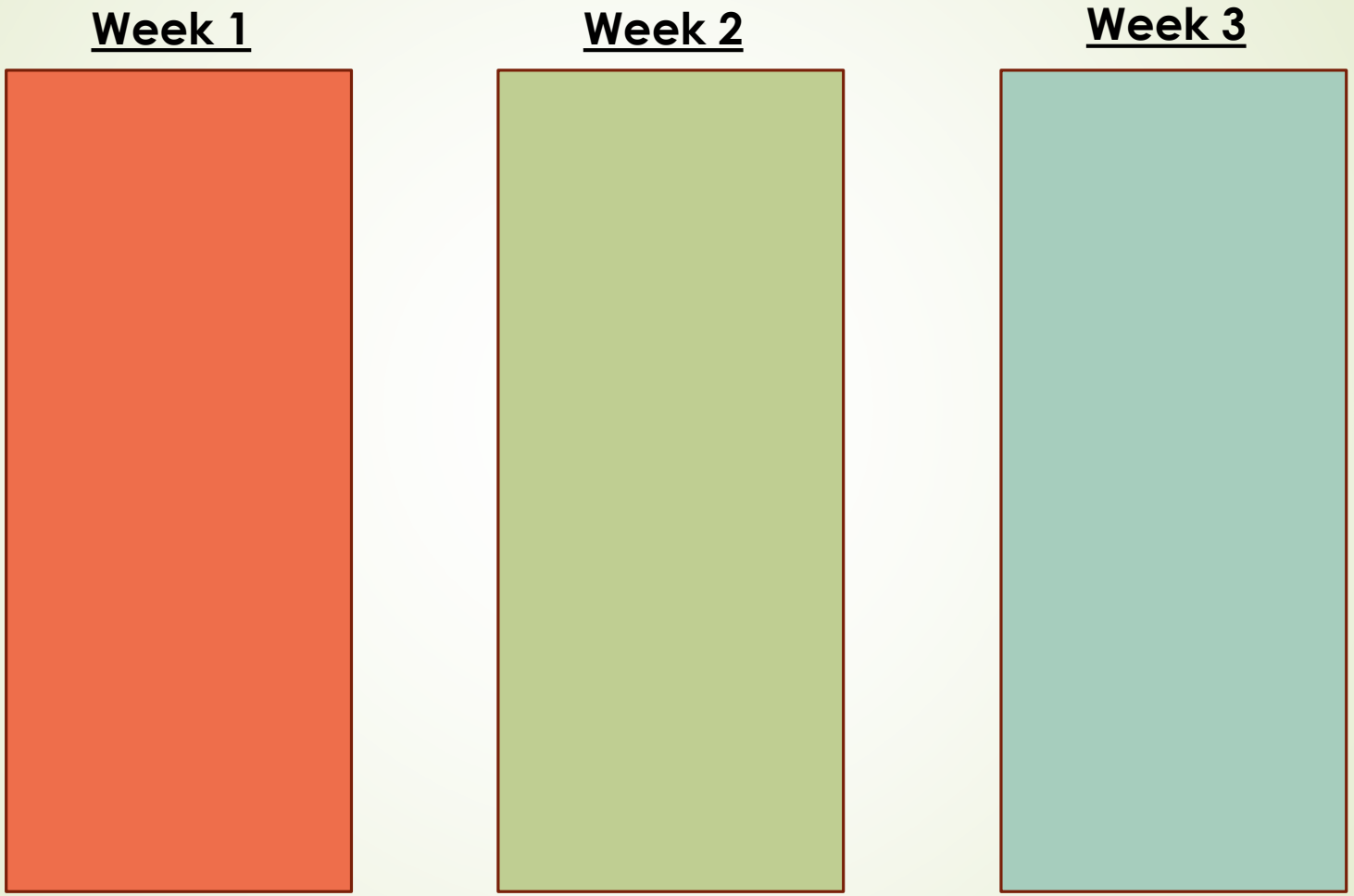
And now, questions and a coffee break!

Let's do an experiment!

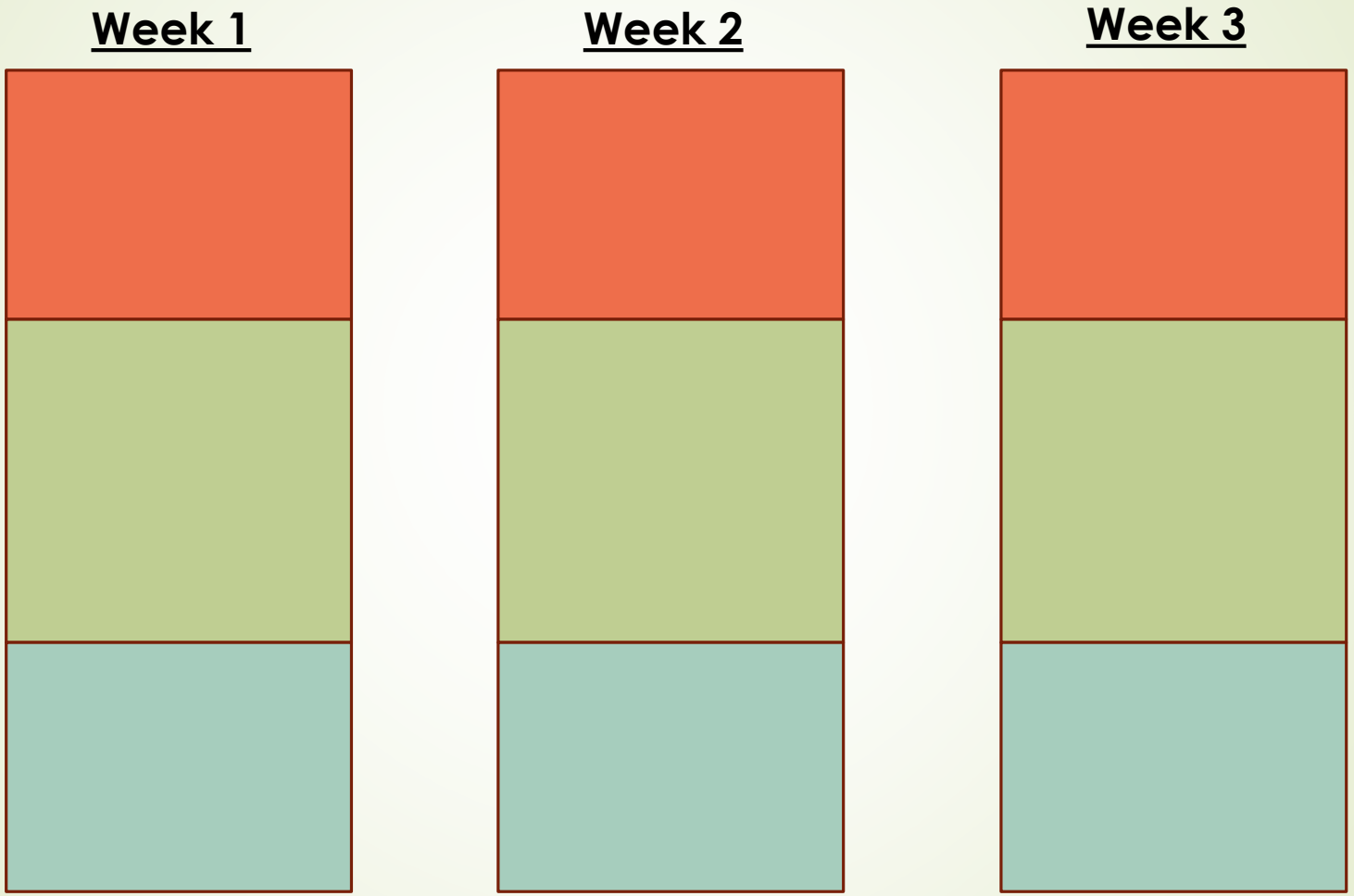
The setup

- 150 individuals
- 50 of each treatment
- Treatment lasts 1 week
- We have 3 incubators/greenhouses/tanks/cages which each hold 50 individuals

Split per week?



Split across weeks?



The twist!

- Let's do the blue treatment in week 1, green treatment in week 2 and red treatment in week 3
  - because … reasons!
- You have 3 undergrads. How should they split the data collection work?
- They are also available for just two days to do the library prep.
- And! You just have 2 lanes per Sequencer available

Let's actually do this in R/RStudio