UNIVERSITY OF HOHENHEIM

Planning and design of experiments

3502-440 Methods of Scientific Working for Crop Science

WS 2024/2025

Prof. Dr. Karl Schmid
Institute of Plant Breeding, Seed Science and Population Genetics
University of Hohenheim

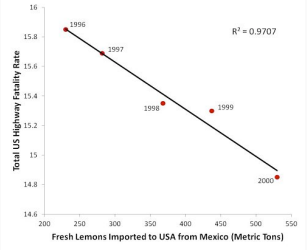
What is an experiment?

Scientific progress:

1. Observations of the phenomenon ⇒ Correlation
   (e.g. Quantitative trait loci (QTL): correlation of genetic markers with a phenotype)
2. Manipulation of the phenomenon ⇒ Causation
   (e.g. Transgenic plants: Transformation with a gene causes a certain phenotype)

Correlation does not imply causation

Higher imports of lemons from Mexico → Fewer people in US are killed on roads



Lee Baker (2017) Truth, Lies and Statistics. Freely available on amazon.com.

- Should the US import more Mexican Lemons?
- Do Mexican lemons kill Americans?
- Post Hoc, Ergo Propter Hoc, meaning 'After this, therefore because of this'?
- In statistics: Post Hoc Fallacy

Experiments need to be designed

... why?

Myth 1:

Just collect data, there will always be a statistical procedure that allows you to analysed the data.

Myth 2:

If you collect enough data there will always be an interesting pattern, and even subtle effects can then be detected.

CHRIS ANDERSON SCIENCE 06.23.08 12:00 PM

THE END OF THEORY: THE DATA DELUGE MAKES THE SCIENTIFIC METHOD OBSOLETE

Illustration: Marian Bantjes

http://www.wired.com/2008/06/pb-theory/

The beginnings of experimental design

James Lind (1716-1794)
Painted by Sir George Chalmers

Painting by Sir George Chalmers

He sailed in 1747 on the HMS Beagle as crew doctor

jameslindlibrary.org

First controlled experiment by James Lind

He took 12 sailors who suffered from scurvy.

Treatments:

1. A quart of cider every day
2. Twenty five gutts (drops) of the elixir vitriol (sulphuric acid) three times a day upon an empty stomach,
3. One half-pint of seawater every day
4. A mixture of garlic, mustard, and horseradish in a lump the size of a nutmeg
5. Two spoonfuls of vinegar three times a day
6. Two oranges and one lemon every day.

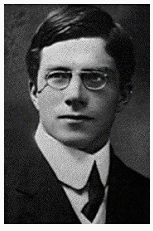
First controlled experiment by James Lind

gave the best result!

Beginning of modern experimental design

Ronald Fisher (1890-1962)

Accepted in 1911 a temporary position at "Rothamsted Research"



Aim: analyse "Classical field experiments"

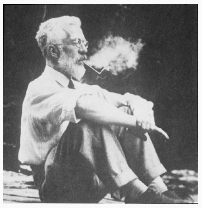
- Broadbalk (started in 1843)
- Alternative Wheat and Fallow (started in 1856)
- Park Grass (started in 1856)
- ...

The design of experiments

R. A. Fisher published in 1935 the book
“The design of experiments”.

Key principles of a good design:

- Randomization
- Replication
- Blocking
- Orthogonality
- Factorial experiments



Randomization

Randomization prevents unwanted side effects due to:

- Grouping of experimental groups
- Position effects
- Effects of the investigator

Replication

Replication allows to estimate the error of a parameter:

- Parameter: e.g. mean or effect size (in a model)
- Error: variance, standard deviation, standard error or residual error

Replication

- What was the replication in the scurvy experiment?

Replication

To find out the replication one needs to go back to the questions!

The scurvy experiment:

Question:	Which treatment heals scurvy?
Answer:	Each treatment level included 2 men.

Blocking

Blocking allows to separate sources of variation that are irrelevant to the results:

- Variation due to a heterogenous experimental field
- Result: a greater precision of the parameter estimate

Heterogeneity in a field trial

Dronefoto: Max Haupt, University of Hohenheim

310 genebank accessions of quinoa

Check varieties (controls)

2 Blocks

Location: Heidfeldhof, Stuttgart

Heterogeneity in a field trial

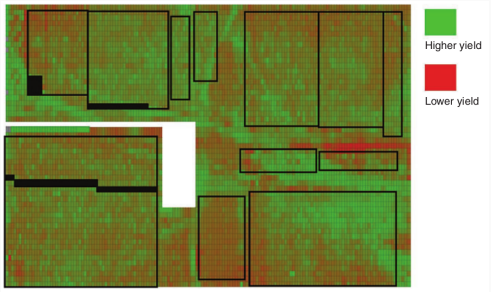


Fig. 5. Heat map of individual plot yield values obtained from an experimental field where multiple experiments were placed into the field to align the experiments with areas of the field with reduced levels of spatial heterogeneity in soil conditions that had been previously characterised. Yield values were obtained from small-plot, combine-harvesting equipment. The experiments were exposed to water-deficit treatments by limiting irrigation during periods when there was no rainfall. The superimposed grid system indicates the individual experimental plots.

Cooper et al. (2014)

Blocking

The scurvy experiment was too small and simple for blocking.

How could we change it in a larger block design?

Orthogonality

At best factors in an experiment are not correlated.

- Each factor can be evaluated independently
- Enables a clearcut partition of variance

Orthogonality

Example of non orthogonal factors:

Experiment to test simultaneously the effect of **Nitrogen (NO^{-3})** and **irrigation**.

Problem:
Nitrogen washes easily out but highly irrigated plants need a lot of it (high growth rate)

Design can be changed to make factors orthogonal.

Factorial design to solve this problem

A factorial design allows for all the above advancements:

- Simultaneaous testing of several factors in same experiment
- **Allows testing of interactions**
  $\Rightarrow$ nitrogen x irrigation
- All subjects in all treatment conditions

	A	B	C	
1	-	-	-	Y1
2	+	-	-	Y2
3	-	+	-	Y3
4	+	+	-	Y4
5	-	-	+	Y5
6	+	-	+	Y6
7	-	+	+	Y7
8	+	+	+	Y8

Disadvantage of factorial design: Needs much space and time: problematic for large experiments.

Theory of experimental design besides Fisher

Abraham Wald
Worked on sequential analyses (e.g. time series) in the 1940s

Abraham Wald

Experimental design after Fischer

For example: Placket-Burman designs

Was developed in 1946 by R. L. Plackett and J. P. Burman, from the British Ministry of Supply

- Not as many replicates as in full factorial design
- Allows for more factors
- Interactions between factors are assumed negligible

Placket-Burman designs

Factors / Experiment	A	B	C	D	E	F	G	H	I	J	K	response
1	+	+	-	+	+	+	-	-	-	+	-	y1
2	-	+	+	-	+	+	+	-	-	-	+	y2
3	+	-	+	+	-	+	+	+	-	-	-	y3
4	-	+	-	+	+	-	+	+	+	-	-	y4
5	-	-	+	-	+	+	-	+	+	+	-	y5
6	-	-	-	+	-	+	+	-	+	+	+	y6
7	+	-	-	-	+	-	+	+	-	+	+	y7
8	+	+	-	-	-	+	-	+	+	-	+	y8
9	+	+	+	-	-	-	+	-	+	+	-	y9
10	-	+	+	+	-	-	-	+	-	+	+	y10
11	+	-	+	+	+	-	-	-	+	-	+	y11
12	-	-	-	-	-	-	-	-	-	-	-	y12
Weightings	0	-10	2	-8	-18	-28	-16	-4	8	-2	10	
Dummy					D2	D1	D3				D4	

Vander Heyden, Y., Nijhuis, A., Smeyers-Verbeke, J., Vandeginste, B.G., & Massart, D.L. (2001) J Pharm Biomed Anal 24, 723-753

Other designs

- Split-plot design
- Alpha-design
- Randomized Balanced Incomplete Block Designs
- Completely Randomized Design
- Augmented block designed
- Latin-square design
- ...

R-Package "agricolae" offers functions to create many of these experimental designs.

Power analysis

Power analysis is an important factor in modern experimental design.

Many statistical software package offer solutions to test the expected power of an experiment.

Power analysis

Procedure:

- Do first small pre-experiment
- Get first estimate of **effect size** and **variance**
- use power-analysis to calculate the approximate sample sizes needed

In **R** the package **pwr** offers functions to calculate the needed group size for a significant p-value

Power analysis

```
library(pwr)
pwr.anova.test(f=0.28,k=4,sig.level=0.05, power=.8)

    Balanced one-way analysis of variance power calculation

             k = 4
             n = 35.75789
             f = 0.28
     sig.level = 0.05
         power = 0.8
```

Note: *n* is number in each group

How to plan an experiment

Identify the best experiment possible:

1. Think a lot about your hypothesis!
2. Make them very clear!
3. Define for each hypothesis separately which evidence you need to falsify it.
4. Combine your needs to a **potential optimal experimental design**

How to plan an experiment

Reality check for the **optimal design**:

1. Perform a **power analysis**
2. Check your spatial capacity (how many samples fit in the greenhouse/on the field?)
3. Check the time planning:
   - How quickly can I process the samples (try it out with a sample and calculate the total needed time)
   - Do I need I involve other people (student helpers, cooperation with colleagues)?
4. Double check the logic of your experiment with your colleagues and supervisor.
5. Double check the required finances with your supervisor.
6. Are all my hypothesis equally important?

Take home messages

Bad design gives bad results
(Also known as 'garbage in garbage out')

In the worst: No results are obtained

Be realistic: Better design a good experiment that allows to test only one hypothesis than a bad experiment that tries to test many hypotheses!

Further reading

- Fisher, RA. 1935. The Design of Experiments. Oxford: Oliver & Boyd.
- Lehmann, EL. 1993. The Fischer, Newman−Pearson Theories of Testing Hypotheses: One Theory or Two? 88: 1242−49.
- Morgan, R W. 1966. How to Design an Experiment. Annals of physical medicine 8(5): 168−71.
- Pukelsheim, F. 2006. Optimal Design of Experiments. New York: John Wiley & Sons, Ltd.

Review questions

1. What is a controlled experiment?
2. Which improvements to experimental design were proposed by R. A. Fischer and why?
3. What is the challenge of a time-series analyses?
4. What is the advantage of a full factorial experimental design?

References i

Cooper, M., Messina, C. D., Podlich, D., Totir, L. R., Baumgarten, A., Hausmann, N. J., Wright, D., and Graham, G. (2014). Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. *Crop and Pasture Science*, 65(4):311.