R03 - Experimental Design

HCI/PSYCH 522
Iowa State University

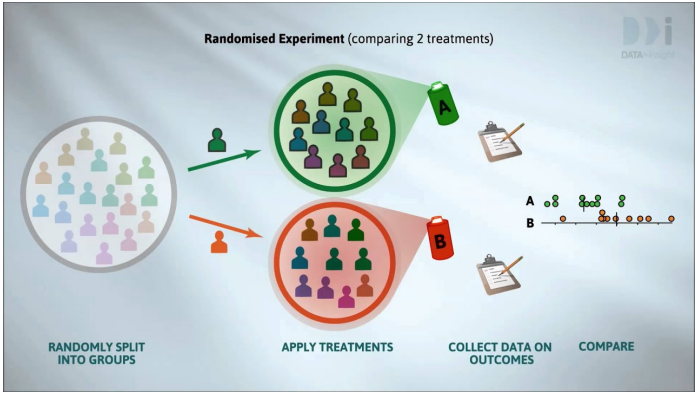
March 10, 2022

Overview

- Terminology
  - Experimental unit
  - Observational unit
  - Factor
    - Treatment
    - Block
- Designs for one treatment variable
  - Completely Randomized Design (CRD)
  - Randomized Complete Block Design (RCBD)
    - Paired design
  - Crossover design
- Designs for multiple treatment variables
  - One-at-a-time
  - Factorial
  - Split-plot design

Randomized Experiments

https://www.youtube.com/watch?v=bi-LNLrFYcQ



Randomized experiments lead to causal inference.

Experimental unit

https://en.wikipedia.org/wiki/Glossary_of_experimental_design

Definition

The experimental unit is the entity to which a specific treatment combination is applied.

Example experimental units:

- Person
- Classroom
- Device

Example treatment combinations:

- Chatbot vs no chatbot
- Working remotely vs going in to work
- Interior design app (IKEA vs yours) & Mode used (online vs downloaded)

Observational unit

https://webspace.maths.qmul.ac.uk/r.a.bailey/Histop/obsunit.html

Definition

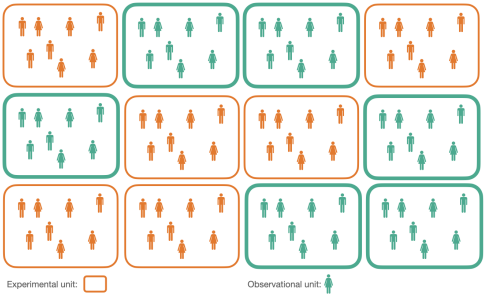
The observational units are what you take measurements on. In many experiments [but not all] they are the same as the experimental units.

Examples of observational units:

- Person
- Person (within a classroom)
- Device

Example experiment

Goal: Assess quality of virtual learning by randomly assigning classrooms to be virtual or in-person and measuring student performance on the final exam.



Simplest approach to analysis is to summarize data within each experimental unit, e.g. average final exam score for a classroom.

Factor

https://en.wikipedia.org/wiki/Glossary_of_experimental_design

Definition

A factor is an variable that an investigator manipulates to cause a corresponding change in the output. Specific values of a factor are called levels. A factor that cannot be assigned by the researcher is a nuisance factor and can often be addressed through blocking.

Example factors:

- Chatbot availability
- Type of instruction
- TA support availability

Example nuisance factors:

- Major
- Classroom
- Age?
- Gender?

Treatment

https://en.wikipedia.org/wiki/Glossary_of_experimental_design

Definition

A treatment (combination) is a specific combination of factor levels whose effect is to be compared with other treatments.

Example treatments:

- Chatbot is available
- Chatbot is not available
- Virtual instruction with no TA support
- In-person instruction with no TA support
- Virtual instruction with TA support
- In-person instruction with TA support

Blocking

Definition

Blocking is the arranging of experimental units into groups (or blocks) that are similar to each other.

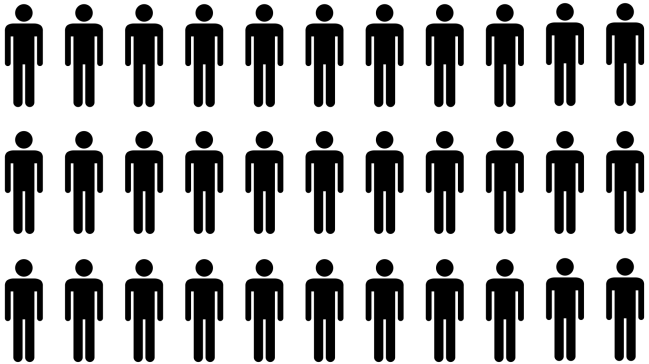
Examples of blocking:

- Age groups
- Gender
- Classrooms

Completely randomized design (CRD)

Definition

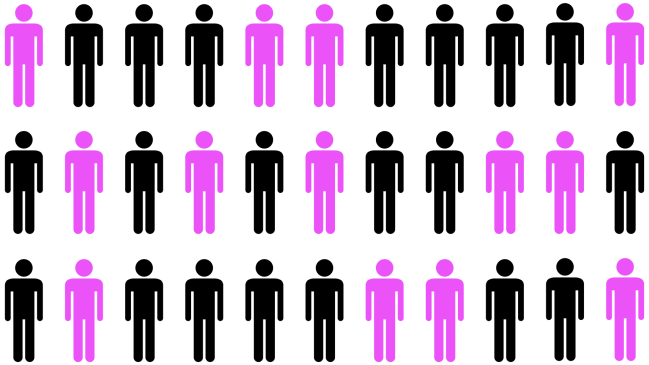
In a completely randomized design, treatments are assigned to experimental units at random.



Completely randomized design (CRD)

Definition

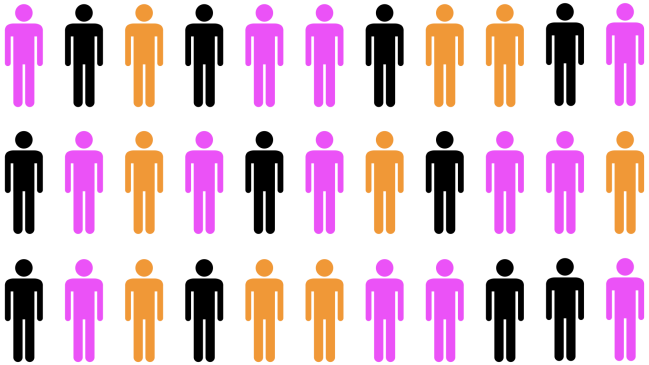
In a completely randomized design, treatments are assigned to experimental units at random.



Completely randomized design (CRD)

Definition

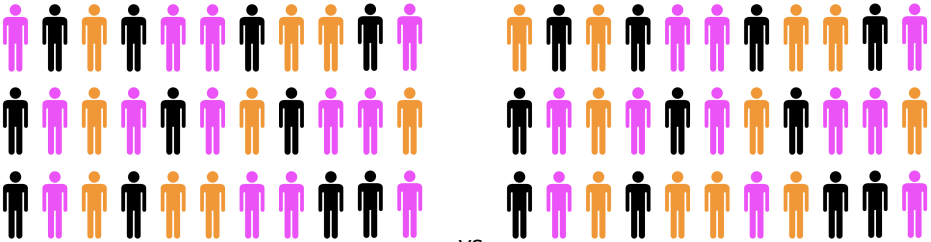
In a completely randomized design, treatments are assigned to experimental units at random.



Balanced

Definition

A balanced experiment contains the same number of observations for each treatment combination.



CRD - Analysis

Appropriate analyses for a completely randomized design:

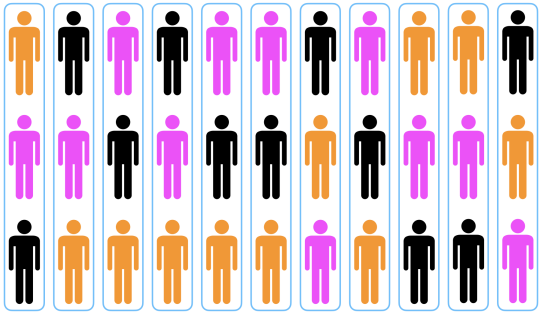
- Comparison of two (or more) probabilities
- Comparison of two (or more) means
- Regression
  - Simple linear regression
  - Categorical independent variable

```
lm(Dependent ~ Independent, data = MyData)
```

Randomized complete block design

Definition

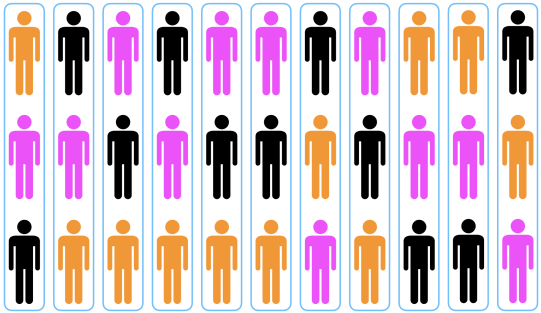
In a randomized complete block design (RCBD), the experimental units are blocked (similar units within a block) and then within each block the experimental units are randomly assigned treatments.



Complete

Definition

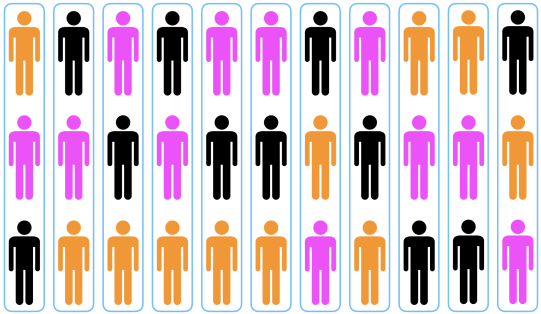
A design is complete if every treatment combination exists in every block.



Replication

Definition

An RCBD is unreplicated if each treatment combination exists in each block at most once. An RCBD is replicated if each treatment combination exists in each block more than once.



Unreplicated RCBD - Analysis

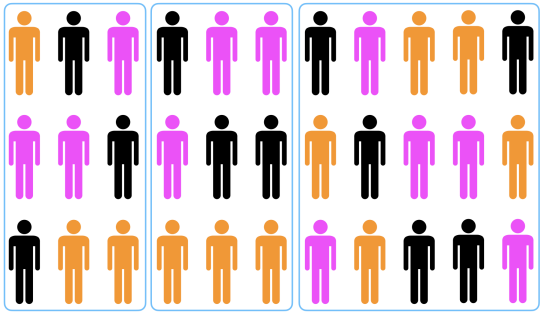
For an RCBD, we need to include both the treatment factors and the blocking (nuisance) factors in the analysis. For an unreplicated RCBD, we fit the main effects regression model.

```
lm(Dependent ~ Block + Treatment, data = MyData)
```

Randomized complete block design

Definition

In a randomized complete block design (RCBD), the experimental units are blocked (similar units within a block) and then within each block the experimental units are randomly assigned treatments.



Replicated RCBD - Analysis

For an RCBD, we need to include both the treatment factors and the blocking (nuisance) factors in the analysis. For a replicated RCBD, we can fit two possible models:

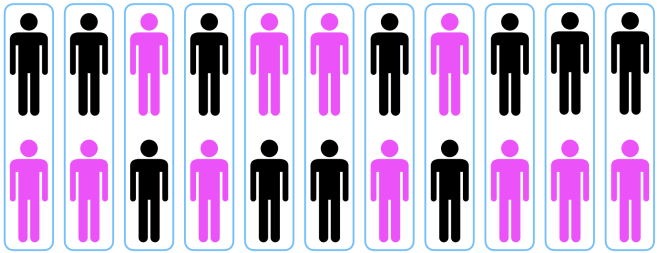
- main effects model or
- a model that include the interaction.

```
# Main effects
lm(Dependent ~ Block + Treatment,                    data = MyData)

# Interaction
lm(Dependent ~ Block + Treatment + Block:Treatment, data = MyData) # or
lm(Dependent ~ Block * Treatment,                    data = MyData)
```

Paired experiment

Definition

A paired experiment is an RCBD with only 2 treatments and a block size of 2.



Paired experiment analysis

Although a paired experiment can be analyzed using the unreplicated RCBD approach, a simpler analysis is available.
For each block, compute a difference (or ratio) of one treatment combination versus the other.
Then analyze that difference (or ratio) using one sample approaches

- Estimating one probability
- Estimating one mean

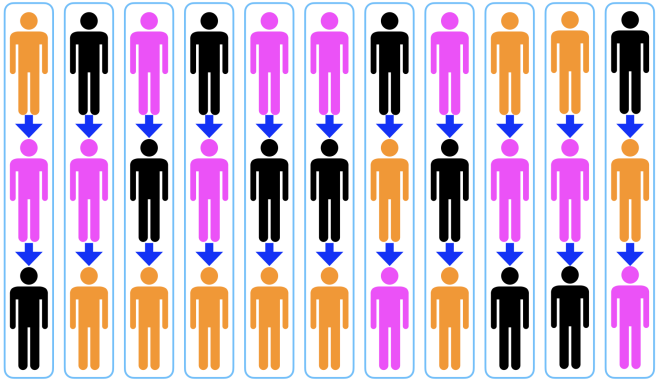
R code:

```
lm(Difference ~ 1, data = MyData)
```

Crossover experiments

Definition

A crossover experiment is a design where the treatment combinations are applied sequentially to the same experimental unit.



Crossover experiments

Benefits of crossover experiments

- Reduce nuisance factor effects
- Efficient use sample size

Limitations of crossover experiments

- Order effects
- Carry-over effects (e.g. learning)

Suggestions

- Try to reduce carry-over effects by using a wash-out period
- Get statistical help to analyze these experiments

Multiple treatment variables

In many (all?) situations, you will have multiple treatment variables of interest.

For example,

- Classroom efficacy study
  - Virtual vs in-person
  - TA hours provided (0, 10, 20 hrs/week)
- Virtual reality interior design app
  - Developer: IKEA vs yours
  - Usage mode: online vs download

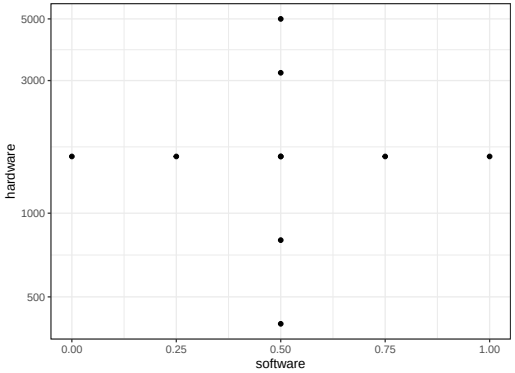
One option is to combine treatment variables into a single compound treatment variable. This is the recommended approach if the design is incomplete.

One-at-a-time design

A common (non-ideal) approach is to investigate one of the treatment variables at a time and hold all other treatment variables constant.

For example,

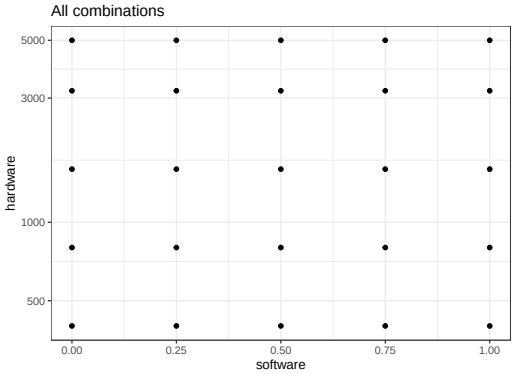
- Mouse accuracy
  - Mouse sensitivity (hardware)
  - Mouse sensitivity (software)

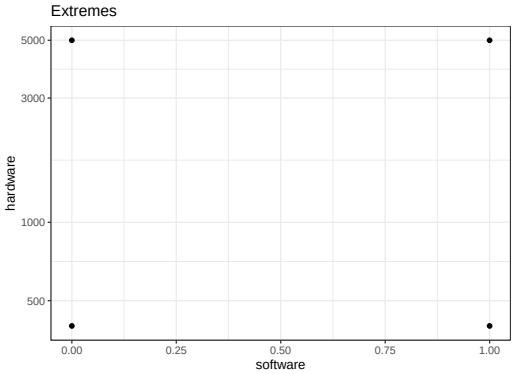


Factorial design

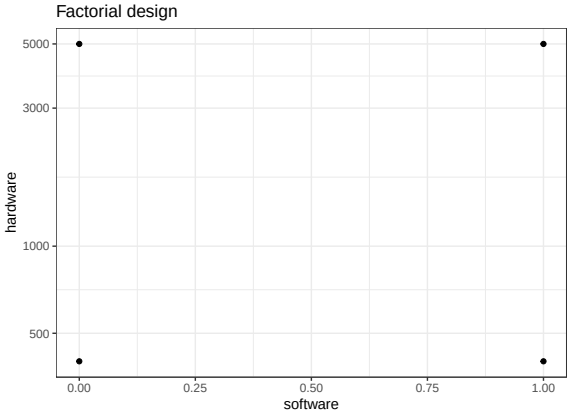
Definition

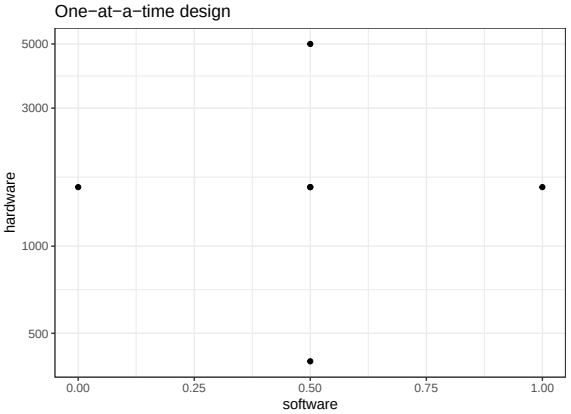
A factorial or (fully crossed) design uses all combinations of all factors in the design.





Comparison: Factorial vs One-at-a-time





Analysis

The analysis for either of these designs is similar to the RCBD design:

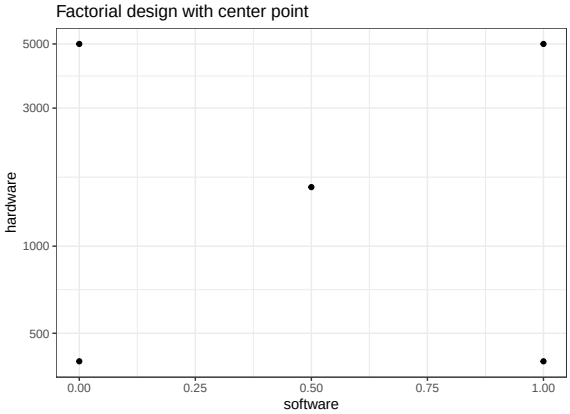
```
# Main effects model
lm(Dependent ~ Factor1 + Factor2, data = MyData)

# Interaction model
lm(Dependent ~ Factor1 * Factor2, data = MyData)
```

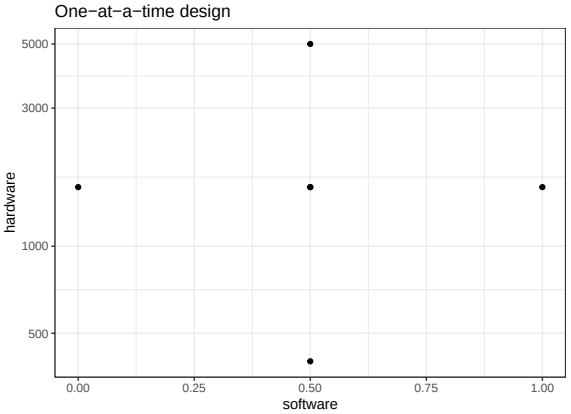
If the factors are continuous (like in the example), you may want to treat them as categorical.

```
MyData <- MyData %>%
  mutate(Factor1 = as.factor(Factor1),
         Factor2 = as.factor(Factor2))
```

Then you can continue with the analyses above.

Adding a center point





Analyses

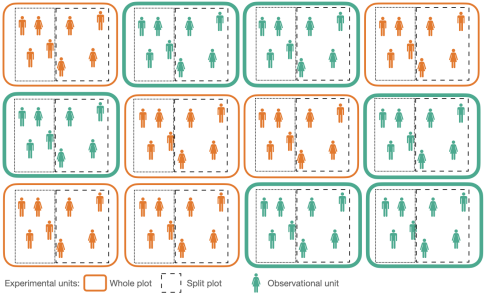
The analyses for designs with a center point, will fit a quadratic curve. Thus factors must be quantitative.

```
# Often we will pre-scale the factors
MyData <- MyData %>%
  mutate(F1 = scale(Factor1),
         F2 = scale(Factor2))

# Response surface model
lm(Dependent ~ F1 * F2 + I(F1^2) + I(F2^2), data = MyData)
```

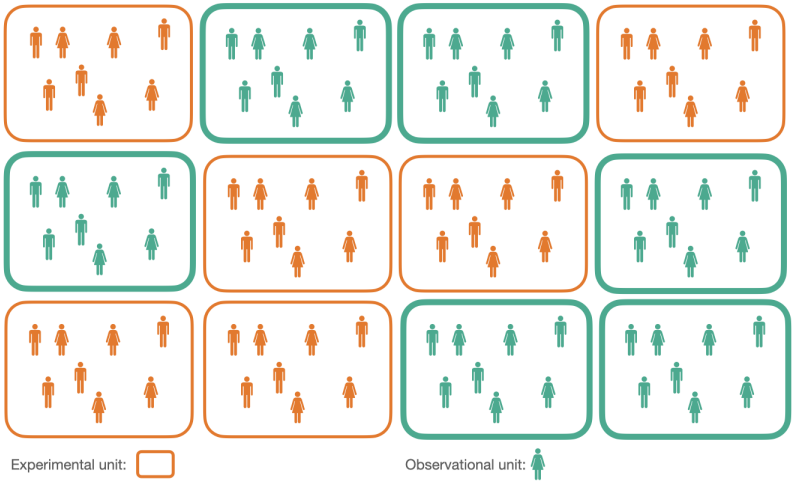
Split-plot experiments

Definition

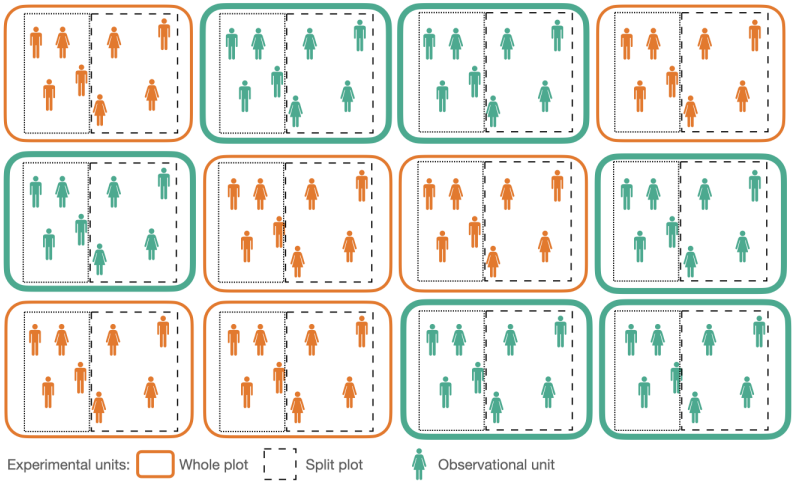
A split-plot experiment occurs when there are (at least) two different treatment factors and there are (at least) two different levels of experimental units.



Split-plot quasi-experiment



Split-split-plot quasi-experiment



Split-plot analysis in R

```
library("lme4") # install.packages("lme4")

lmer(Dependent ~ Factor1 * Factor2 + (1|WholePlot), data = MyData)
```

Summary

Think carefully!