Study design

- Generalizing
  - How do I want this to generalize?
  - What population to generalize to?
  - What is the scope of inference?
- Generalization is determined by the design not the analysis
- Study design is best done before data collection
  - simulation!

Study design

- Observational design
  - focus: sampling
  - estimation, prediction, weaker causal inference
- Experimental design
  - manipulative
  - causal inference

To find out what happens when you change something, it is necessary to change it

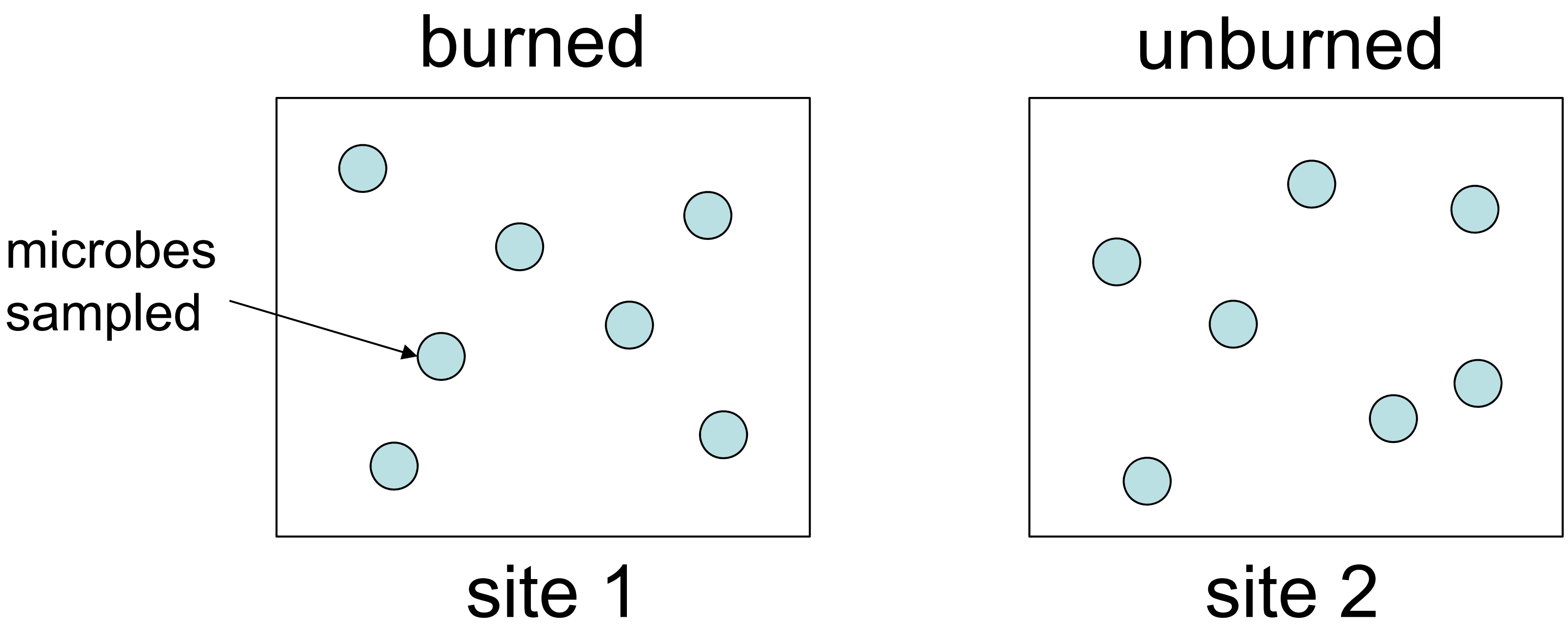
Box, Hunter, and Hunter (1978)

Design fundamentals

- Identify a population of inference: scope
- Identify sample or experimental unit
- Confounding — main issue
- Replication
- Randomization
- Control

main remedies

Confounding examples



burn and site are confounded

Confounding examples

Process all of treatment 1

before lunch

Process all of treatment 2

after lunch

What's wrong?

Confounding examples

Process all of treatment 1

before lunch

time 1
environment 1?

Process all of treatment 2

after lunch

time 2
environment 2?

treatment and time are confounded

Confounding examples

Put all of treatment 1

left side of bench

Put all of treatment 2

right side of bench

What's wrong?

Confounding examples

Put all of treatment 1

left side of bench

space 1
environment 1?

Put all of treatment 2

right side of bench

space 2
environment 2?

treatment and space are confounded

Replication

- How much replication?
  - depends on effect size and variance
  - rule of thumb:

    < 20 d.f. is treacherous

    > 100 d.f. is good (but unusual)
- Degrees of freedom (d.f.)

  = n – number of parameters
- Best to simulate designs

Pseudoreplication

- Replicates are grouped
- Grouping = confounding

Randomization

- Fixes confounding by shuffling potential confounders
- Random sampling: easiest inference to population (scope)
- Random assignment: allows causal inference about a treatment

Simple random sample

- Number each individual in the population
- Use a random number generator to draw individuals at random
- Unbiased sample
- Ensures unbiased estimate