Factorial ANOVA

Spring 2023

Outline

- Factors
- Study Design
  - T-test
  - ANOVA
  - Factorial
- Blocking
  - Latin Square

Factors

- Research Questions often interested in more than one independent variable.
- Example
  - Effect of Fertilizer vs. No Fertilizer
  - Effect of different fertilizers: Nitrogen, Phosphorous, no fertilizer
  - Effect of types of fertilizer: Animal manure, synthetic liquid Nitrogen/Phosphorous, slow-release pellets…. Etc.

Factors and levels

- *Factors* are the predictor variables
- Different values of factors are the *levels*
- Examples:

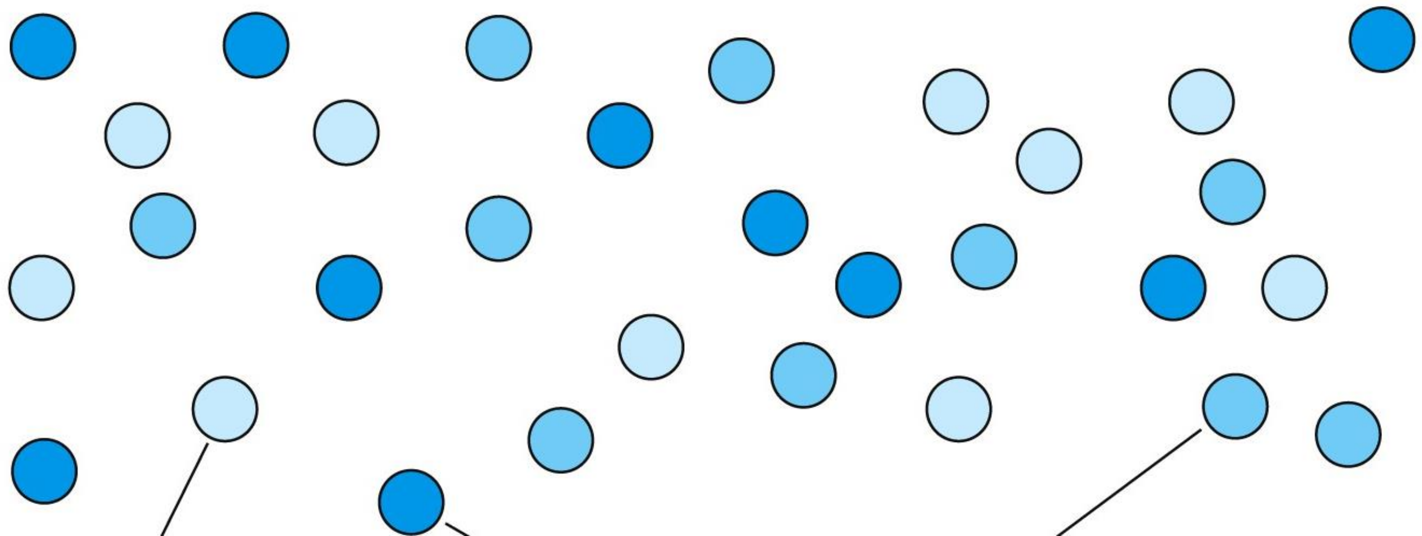
Factors

Levels

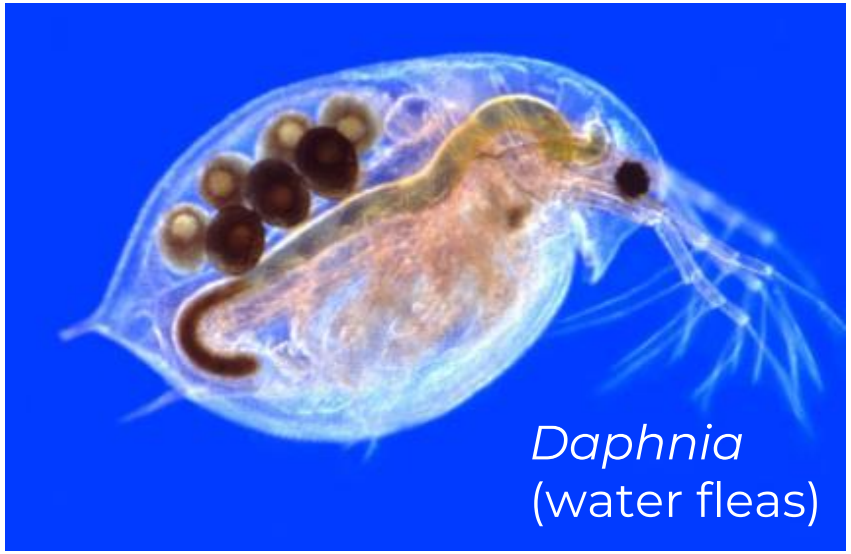
Nitrogen	Irrigation	Species	Predators	Temperature
None	None	*G. scandens*	Control	Ambient
Low	Uniform	*G. Fortis*	Exclusion	+ 1 C
Medium	Supplemental	*G. Fuliginosa*	Addition	+2 C
High	--	*G. Difficilis*	--	+5 C

Basic study design

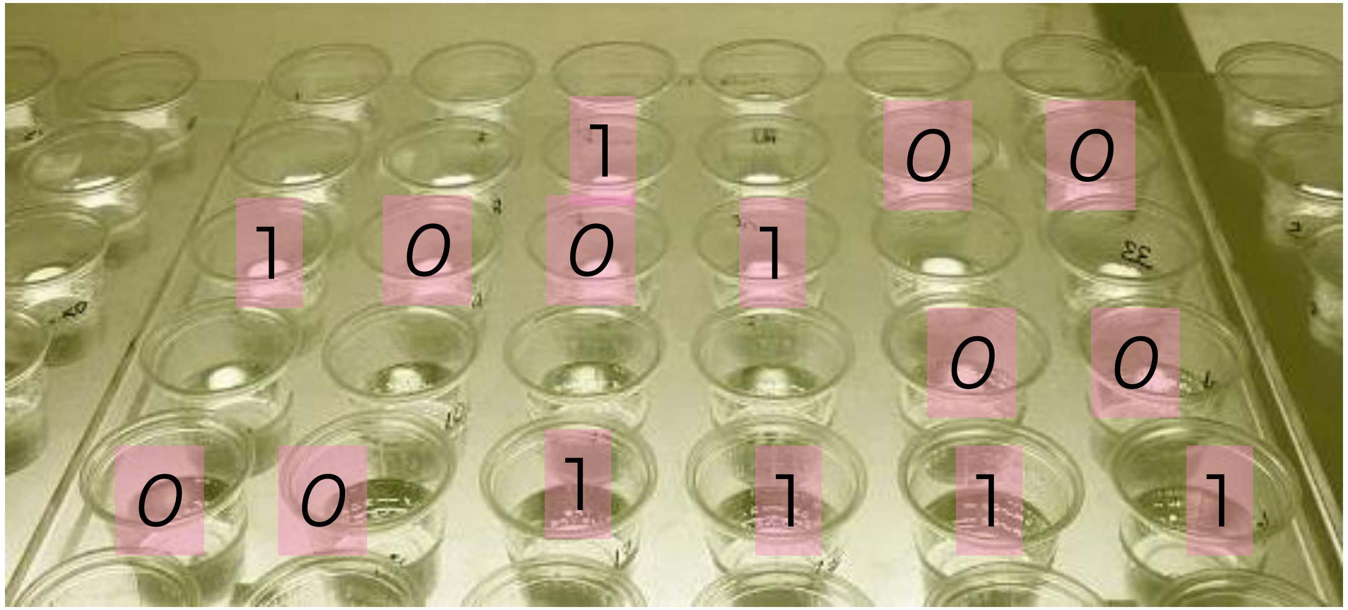
- Minimize (eliminate) all variation except for main factor testing
- Treat all units the same, only modify the factor of interest.
- Control treatments
- Random assignment to treatment level



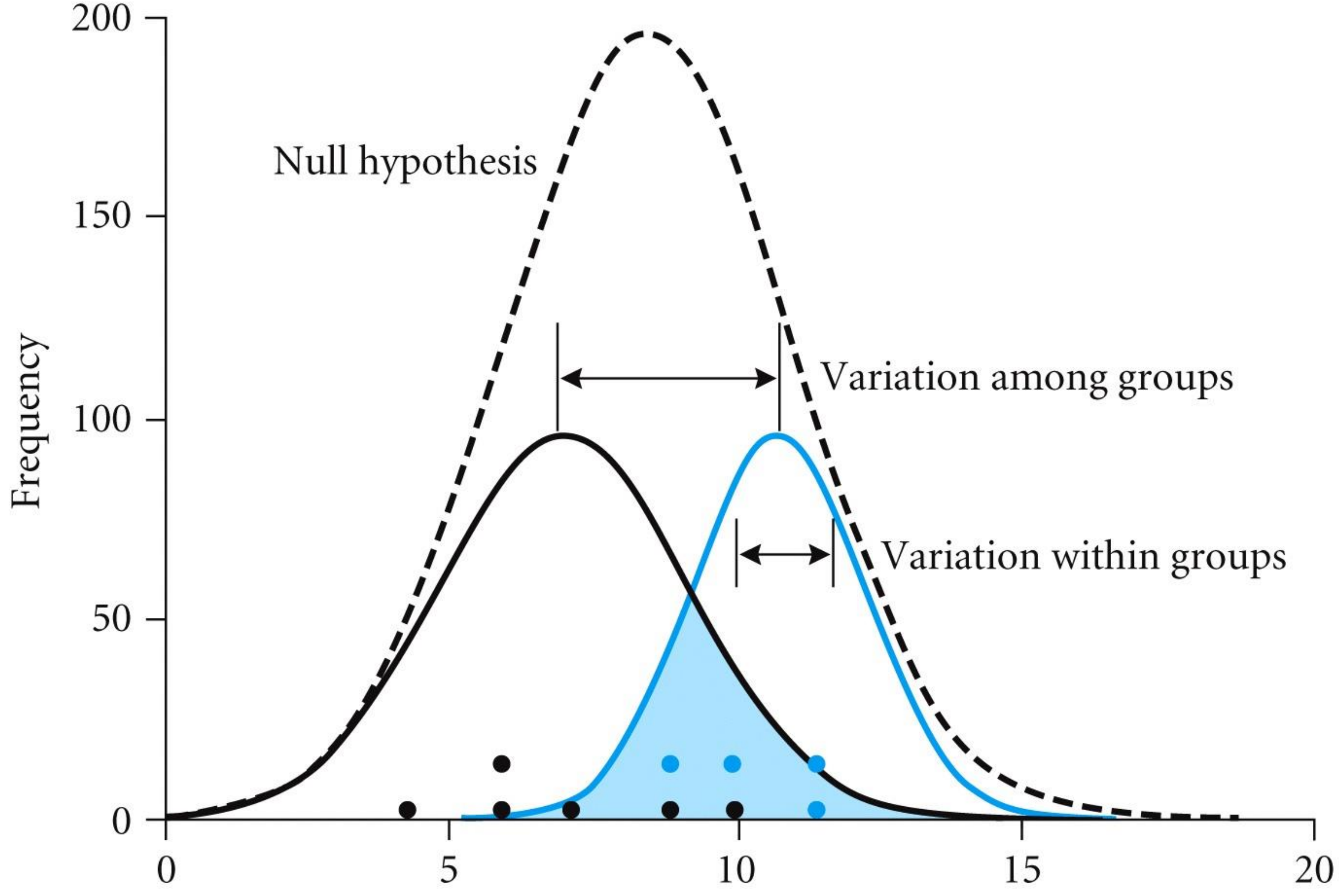
Experimental design: T-test



- One factor, two levels
- Multiple replicates
- Example:
  - Does Competition slow growth of *Daphnia*?
  - Factor: Competition
  - Levels:
    - None (control); 0
    - Competitors; 1
  - Replicates = "cups" randomly assigned to treatment level

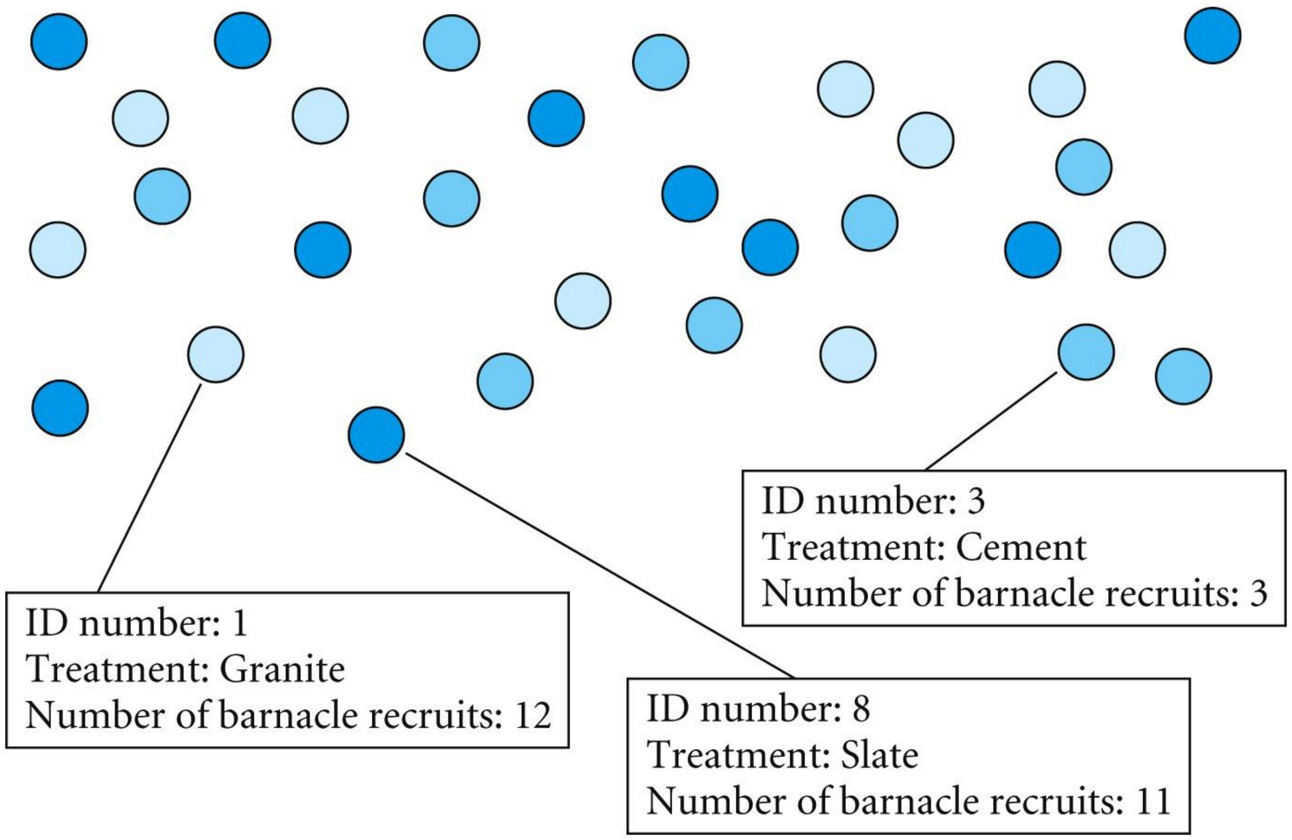


Analysis: are means different?



One-way ANOVA

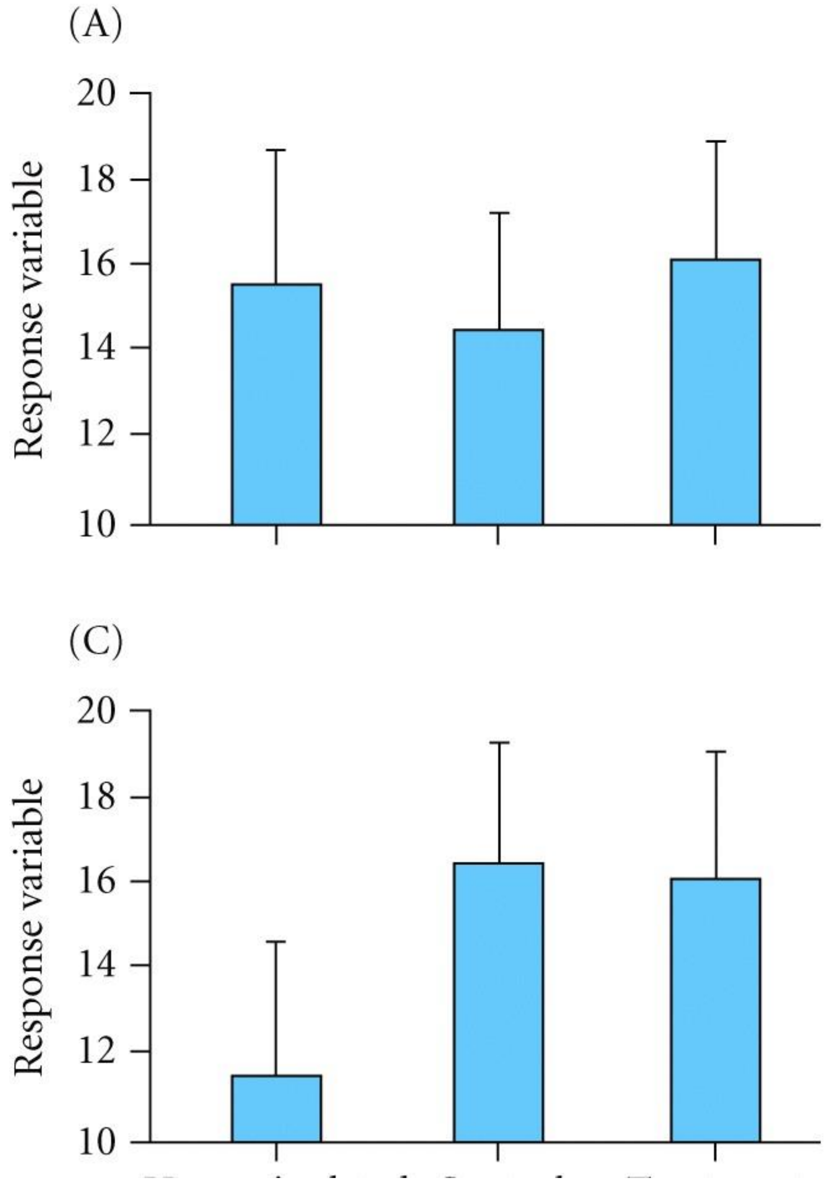
- One factor, 2+ levels
- Multiple replicates
- Example:
  - Does Barnacle recruitment depend on substrate?
  - Factor: Substrate
  - Levels: different shades of blue
  - Replicates = each circle is a “plot” or replicate
    - Requires many more reps



A PRIMER OF ECOLOGICAL STATISTICS 2e, **Figure 7.5 (Part 1)**

One-way analysis

- Model: `y ~ factor`
- All means are equal (top panel)
- Or at least one mean is different (bottom panel)



One-way ANOVA table

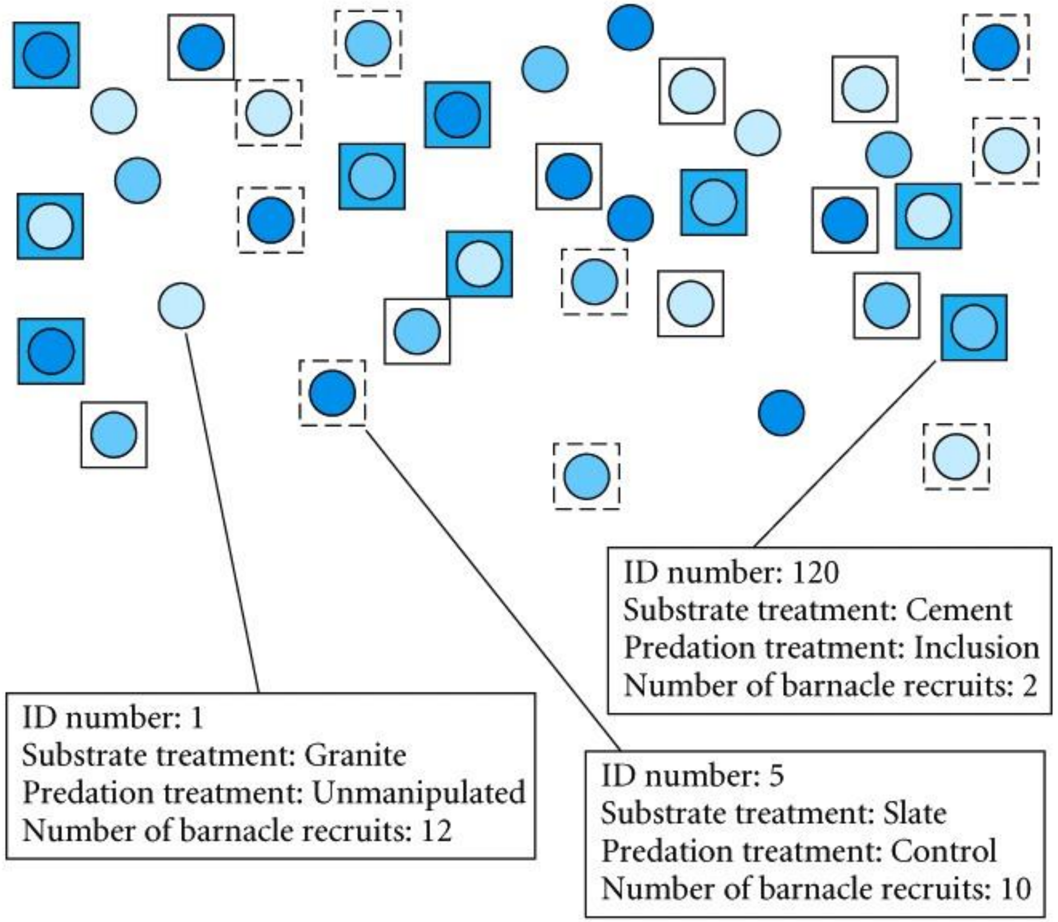
```
> anova(lm(Values ~ Group, dataOneWay))
Analysis of Variance Table

Response: Values
          Df Sum Sq Mean Sq F value    Pr(>F)
Group      1     60  60.000  64.444 5.503e-11 ***
Residuals 58     54   0.931
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- `Group` is the factor of interest
- Low p-value, factor is significant
- At least one mean in the `Group` is different

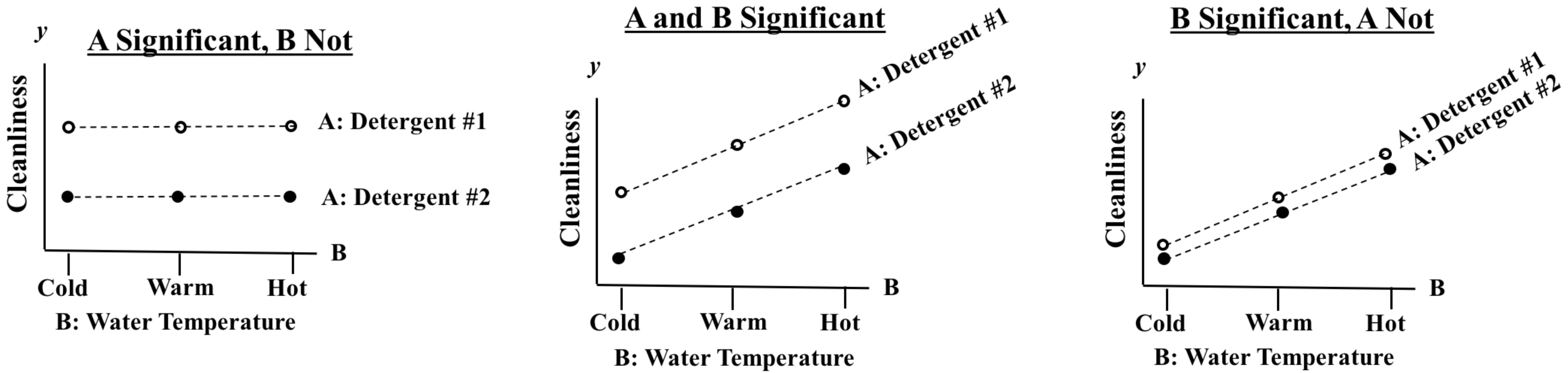
Factorial ANOVA

- Multiple factors, $\geq$ 2 levels
- Does Barnacle recruitment depend on substrate *AND* predation?
- Factors:
  - Substrate: different shades of blue
  - Predation: shape outline
- As factors and levels increase, need many, many more replicates
  - Regression design usually better option (more on this later



Factorial analysis

- Model: `y ~ factor`$_A$ `+ factor`$_B$
- Main effect of A (detergent) and B (temperature) assessed separately
- Effect of one factor while controlling for effect of other factor



Factorial analysis

- Main effect of A (detergent) and B (temperature) assessed separately
- Effect of one factor while controlling for effect of other factor
- Analyze both factors at once, pools variation
- **NOT** the same as running two separate one-way ANOVAs

Two Factor ANOVA table

```
anova(crop_lm)
```

```
## Analysis of Variance Table
##
## Response: yield
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## density     1  5.1217  5.1217 15.3162 0.0001741 ***
## fertilizer  2  6.0680  3.0340  9.0731 0.0002533 ***
## Residuals  92 30.7645  0.3344
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Main effect of A (**`density`**) and B (**`fertilizer`**)
- Both are significant, when controlling for impact of the other
- No information on how they're different, coefficient estimates, etc. (Need **`lm()`** or other analysis for that)

Multiple Factors, Multiple levels

- Factorial analyses are essentially limitless
- However, replicates for more complicated designs quickly become limiting

Multiple Factor ANOVA table

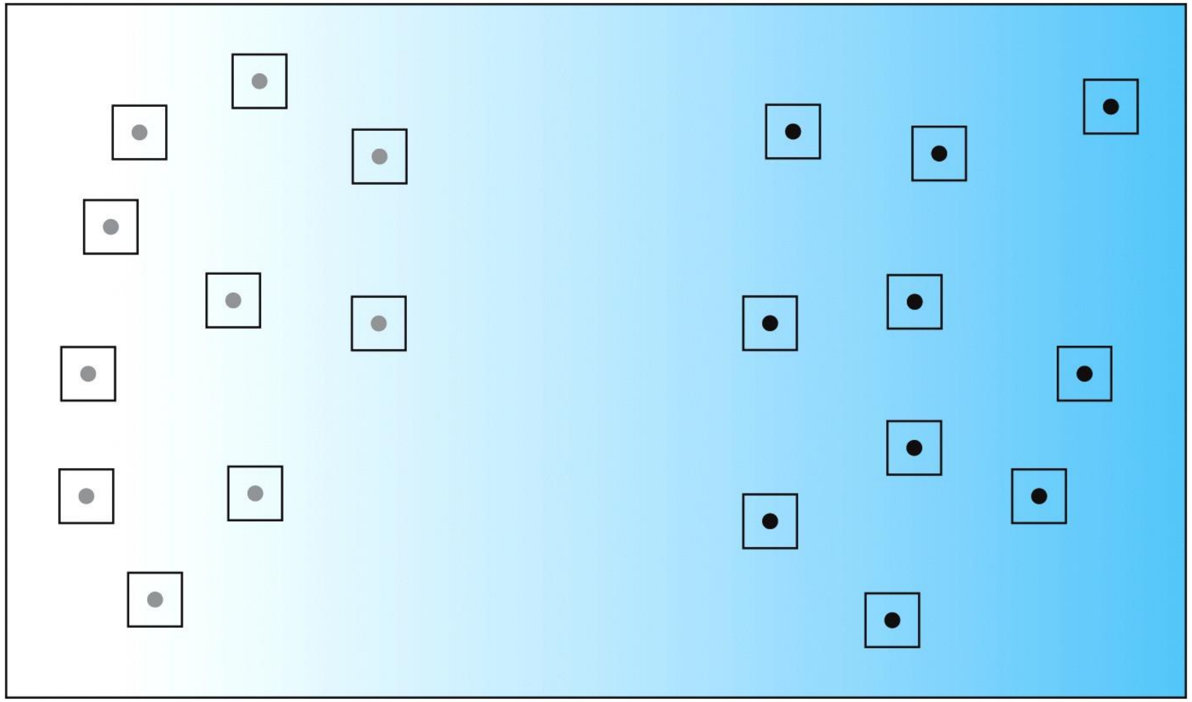
```
## Analysis of Variance Table
##
## Response: yield
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## density     1  5.1217  5.1217 15.2238 0.0001840 ***
## fertilizer  2  6.0680  3.0340  9.0184 0.0002693 ***
## block       2  0.4861  0.2431  0.7225 0.4883291
## Residuals  90 30.2784  0.3364
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Assess the *overall* effect of each factor
- Results are only based on including other factors in model
  - Look at previous 2-factor table (w/o **block**), statistics are slightly different

What about confounding factors?

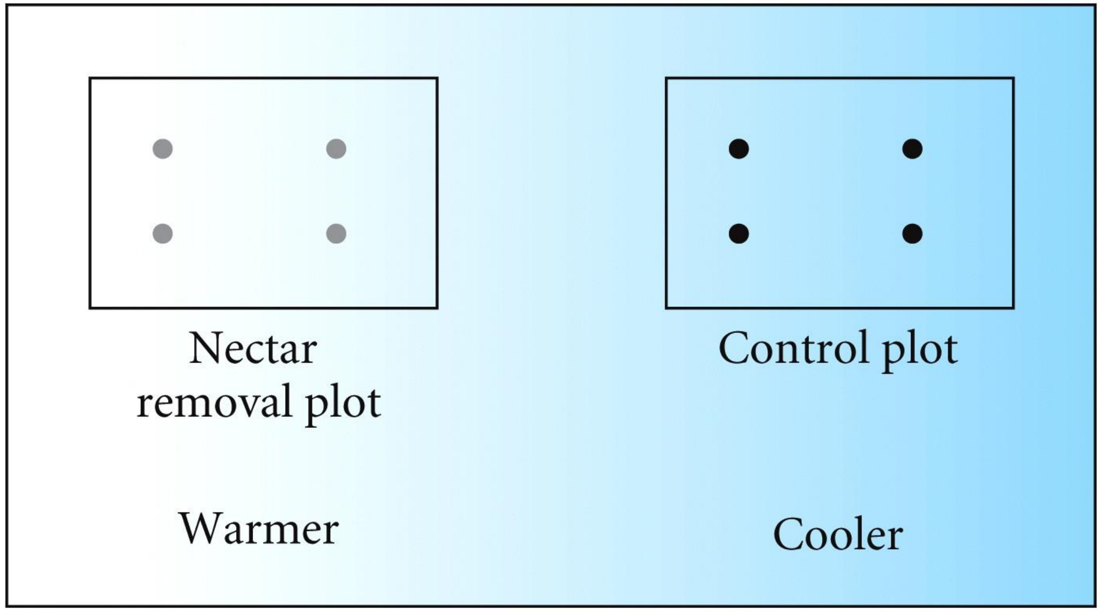
- Factor of interest: what we're testing or manipulating
  - i.e., Fertilizer
- Confounding factors: natural variation across experimental units or plots.
- Example:
  - Plot gradient, dry to wet
  - Elevation gradient
  - Temperature gradient

Confounding Variables



A PRIMER OF ECOLOGICAL STATISTICS 2e, Figure 6.8

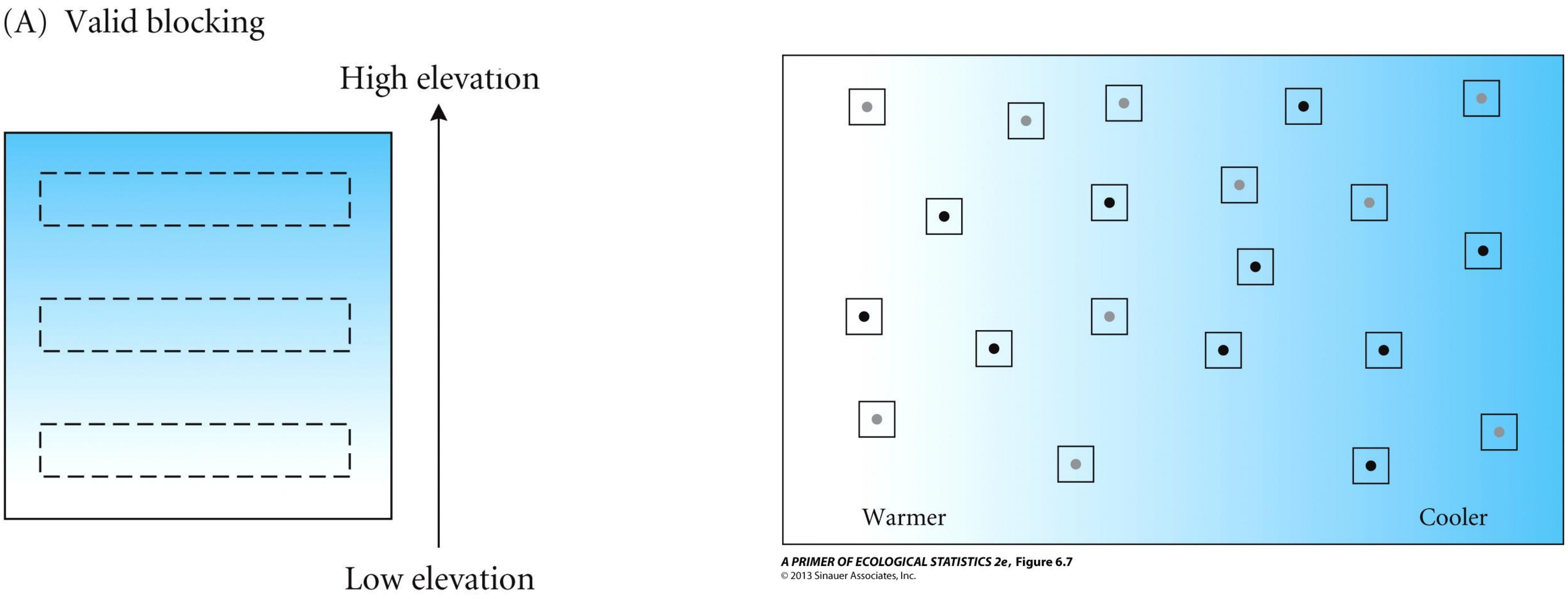
Moisture Gradient



Temp. Gradient

Controlling for confounding factors

- Blocking factors
- Groups organized based on known (or unknown) confounding variables.

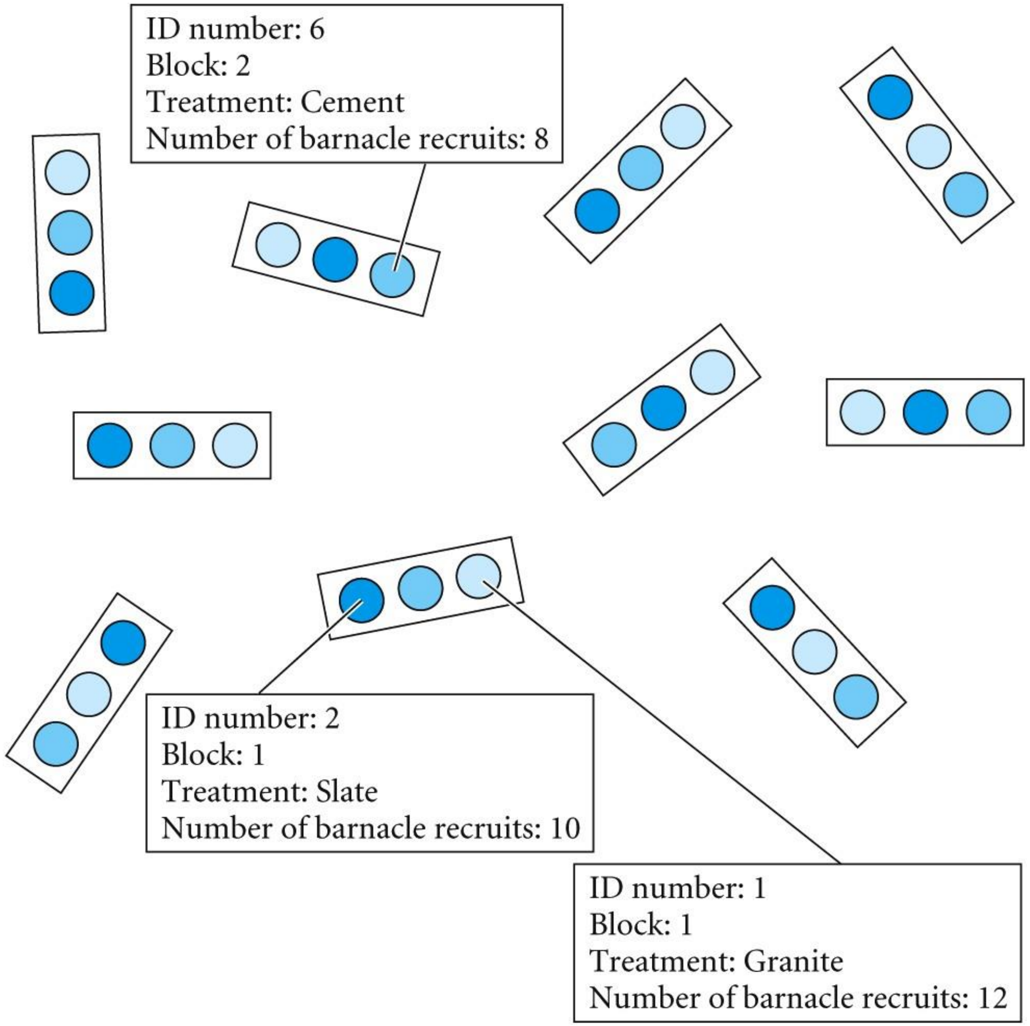


A PRIMER OF ECOLOGICAL STATISTICS 2e, **Figure 6.7**

Controlling for confounding factors

- Randomized Complete Block Design (RCBD)
- Block = rectangle (Right) or Row (Bottom)
- All treatment levels present in each block, each block replicated
- Model: `Y ~ Factor + Block`

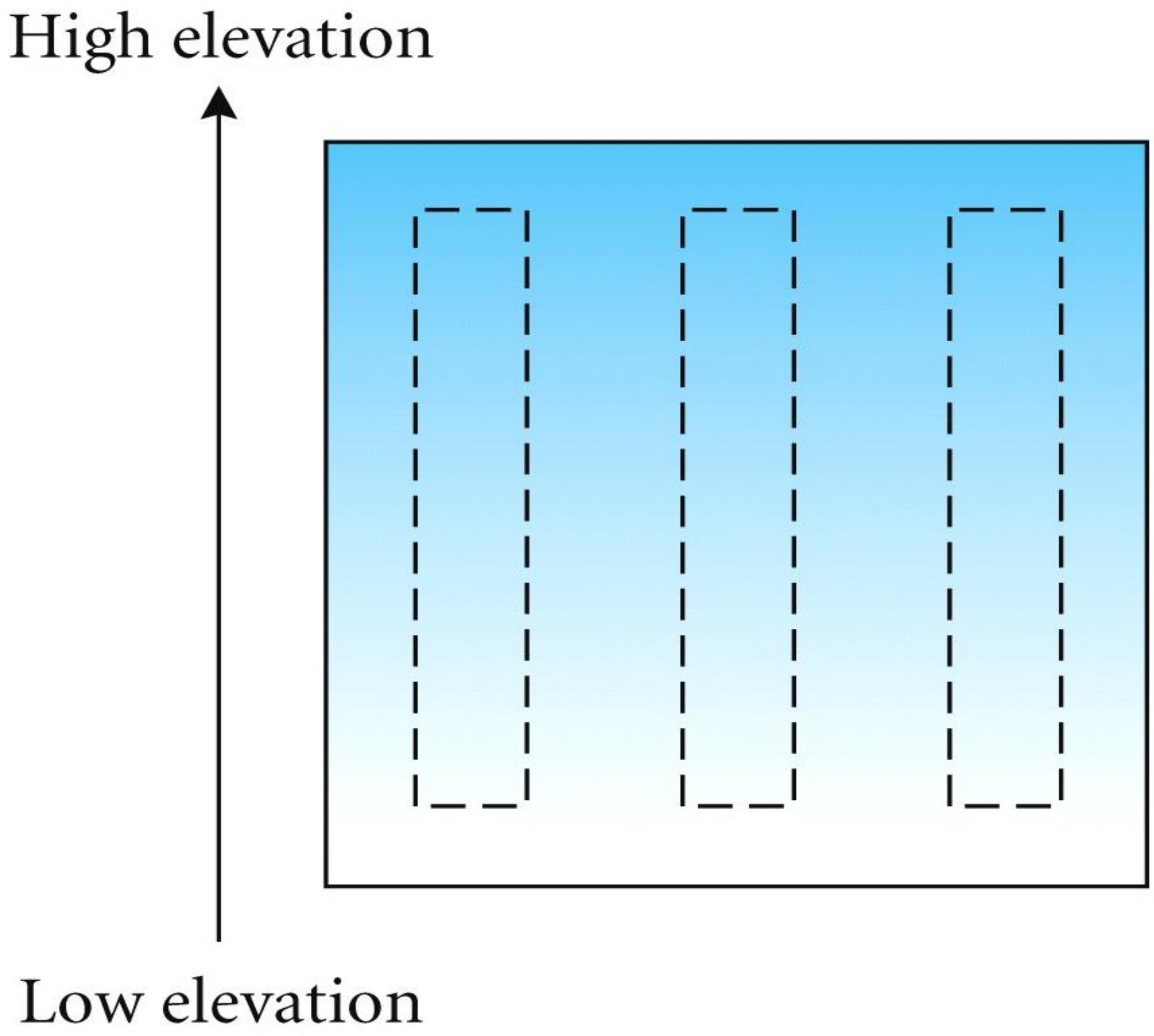
	Rep I	Rep II	Rep III
Block 1	Group A	Group C	Group B
Block 2	Group C	Group B	Group A
Block 3	Group B	Group A	Group C



***A PRIMER OF ECOLOGICAL STATISTICS 2e*, Figure 7.6 (Part 1)**

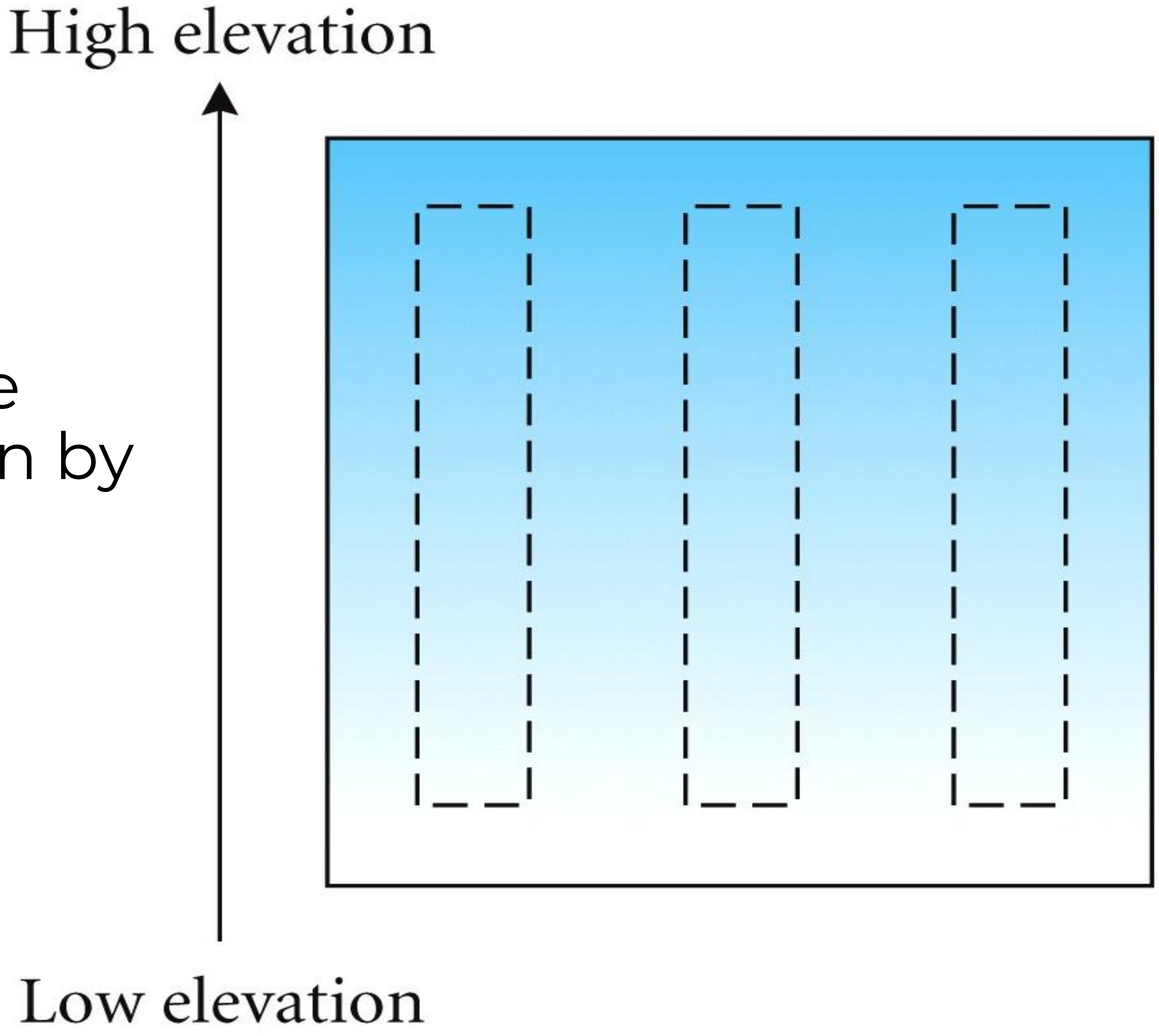
Invalid Blocks

- Why?



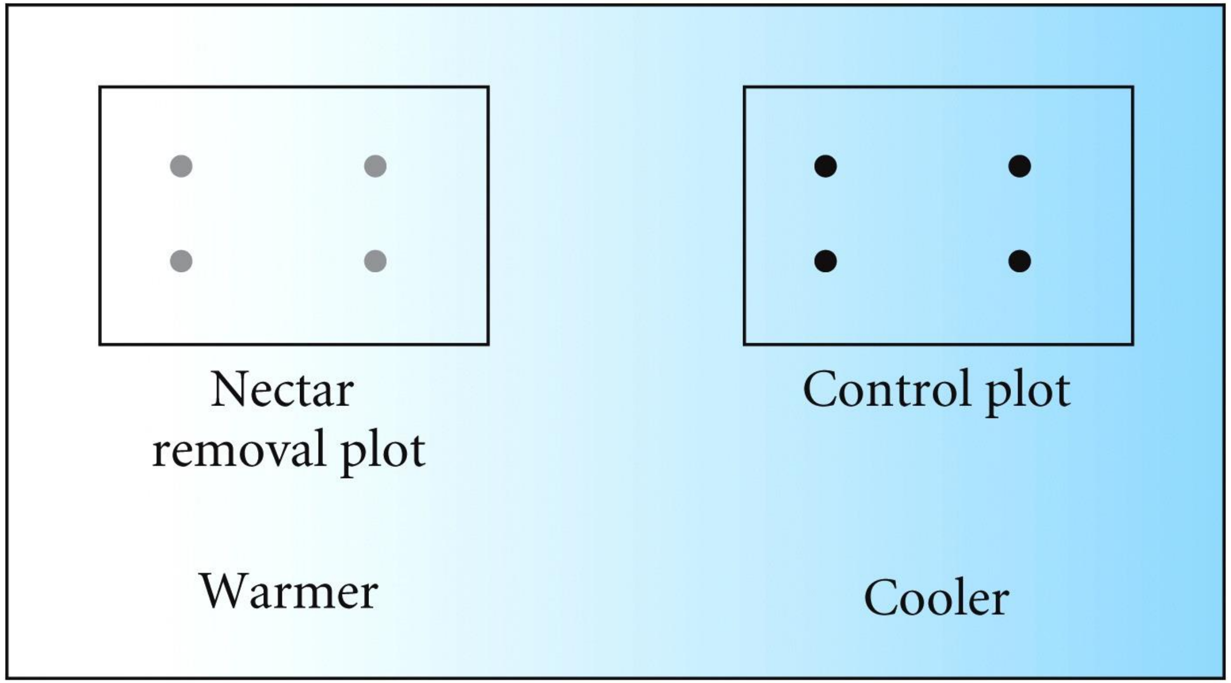
Invalid Blocks

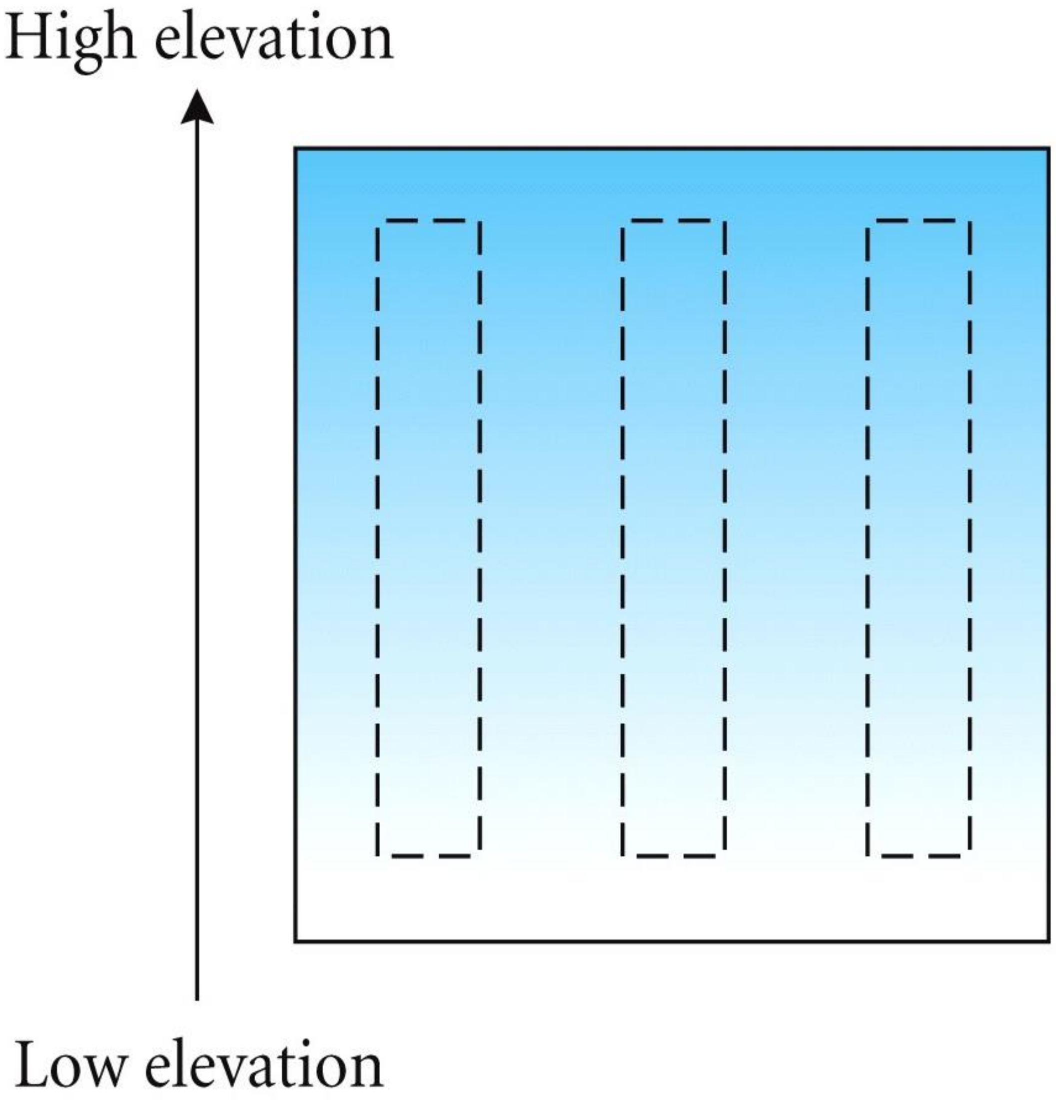
- Why?
- Treatment levels may be grouped at one elevation by chance



Invalid Blocks

- Why?
- Treatment levels may be grouped at one elevation by chance





Special blocked designs

- Double-blocks: Latin Square
- Moisture gradient L to R
- Temp. gradient Top to Bottom
- Each level appears once in each row and column (no replication)
- Powerful design to test 2- or 3-factors, no interaction possible

Rows	Columns			
	1	2	3	4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

4×4 Latin Square Design

Latin square for 3-factors of interest

- Response = Yield
- “Rows” = fertilizer mix
- “Columns” = tillage method
- Treatment = “cells”

Model: `y ~ treatment + row + column`

Latin square ANOVA table

```
## Analysis of Variance Table
##
## Response: l_decrease
##           Df Sum Sq Mean Sq F value   Pr(>F)
## treatment  7 62.904  8.9862 51.5196 < 2e-16 ***
## rowpos_f   7  2.768  0.3955  2.2672 0.04726 *
## colpos_f   7  0.850  0.1214  0.6960 0.67498
## Residuals 42  7.326  0.1744
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Factors: `Treatment***, rowpos_f*, colpos_f` (NS)
- Which row the plot was in did have an effect. `Treatment` also had an effect, when controlling for the effect of row position.