Validity of Experimental Designs and Analyses

Randomization

Avoiding experimenter bias (accuracy)

Don't preferentially assign some types of EUs to some treatments

Replication

Improve precision Reduce $\sigma_r(\hat{\delta})$

Estimate standard errors s^2

Check Assumptions

QQ plot, S/L plot

* using estimates of EUs

Treatments

Positive, negative controls

Manipulation required for causal inference

Confounding factors?

Replication

Replication of **Treatment Levels** “Experimental Unit” or “EU”

Replication of **Treatment Effects** “Replicate”

2)

	Jill	Bob	Amy
T1	Sit	Stand	Sit
T2	Stand	Sit	Stand

# people	# measures	#EU
40	80	80

Structure	Variable	Type	#levels	Replicate **Treatment Levels**	EU **Treatment Effects**
Treatment	Posture	Cat	2	Person	Person:Trial
Design	Person	Cat	40		
	Trial	Cat	2		
	Person:Trial	Cat	80		
Response	Pulse	Num	80		

Replication of **Treatment Levels** is **required** for valid experiments

Replication of **Treatment Effects** is more powerful

- Requires **valid** EUs
- Controls for variation among EUs
- Allows for better generalization

Example 1:

40 pots are planted with pepper plants

1 plant per pot

2 hot and 2 cold growth chambers

10 pots per chamber

2 leaves harvested per plant (pot)

RNA extracted from each leaf

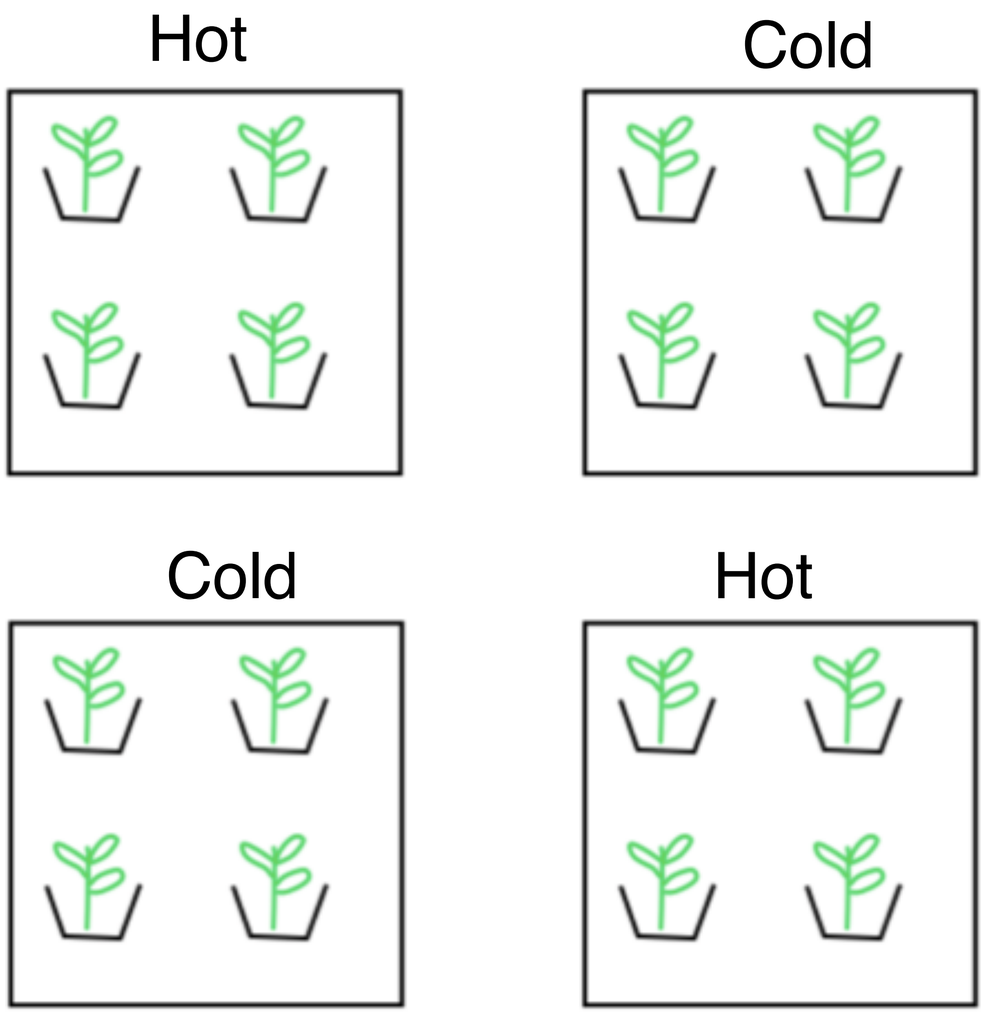
expression of the gene *sp1* measured 3 times per RNA sample

What is the Experimental Unit?

Is there a Replicate of the treatment effect?

Is the Treatment (temperature) confounded with any other factor?

Why isn’t the EU the Plant?



Replicate of Treatment:	Chamber
Replicate of Chamber:	Plant
Replicate of Plant:	Leaf

The **smallest** unit of experimental material to which a **single treatment** (or treatment combination) is assigned by the experimenter and which is dealt with **independently** of other such systems **under that treatment** at **all stages in the experiment** at which important variation may enter.

Kozlov and Hurlbert 2006

EUs should be **randomly selected** from a **reference population**

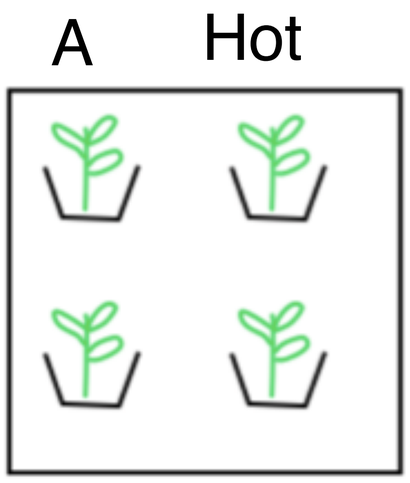
Each EU is **equally likely** to be assigned each treatment

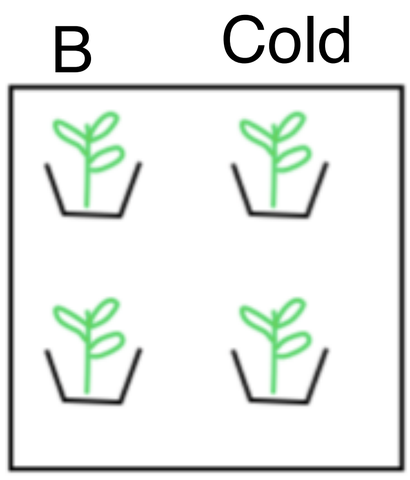
EUs shouldn’t **interfere** with each other

Treatments are applied **independently** to each EU

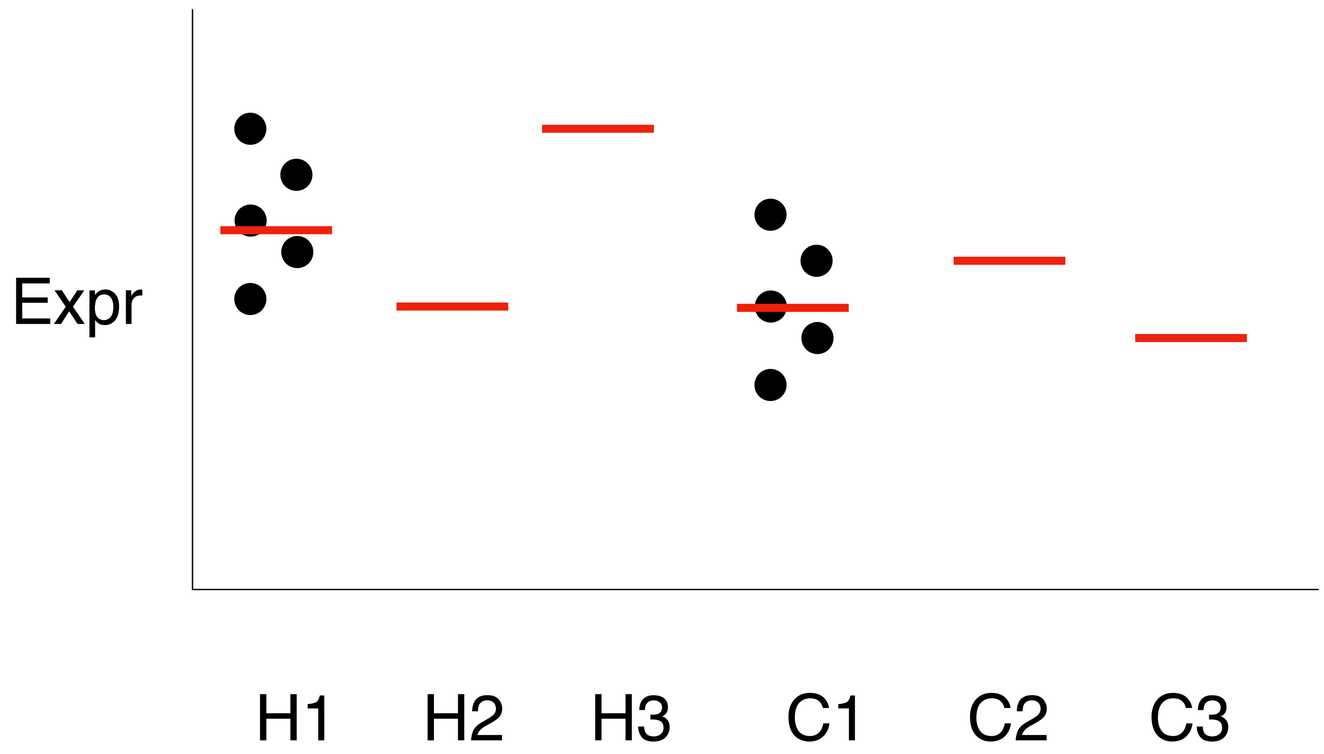
EUs are interspersed **temporally and spatially**

Why isn’t the EU the Plant?





A: Hot



Say you forgot to set the temperature…

Would you detect a difference between Hot and Cold?

Add 100 plants. Would you now?

Add 10 chambers. Would you now?

Potential sources of confusion

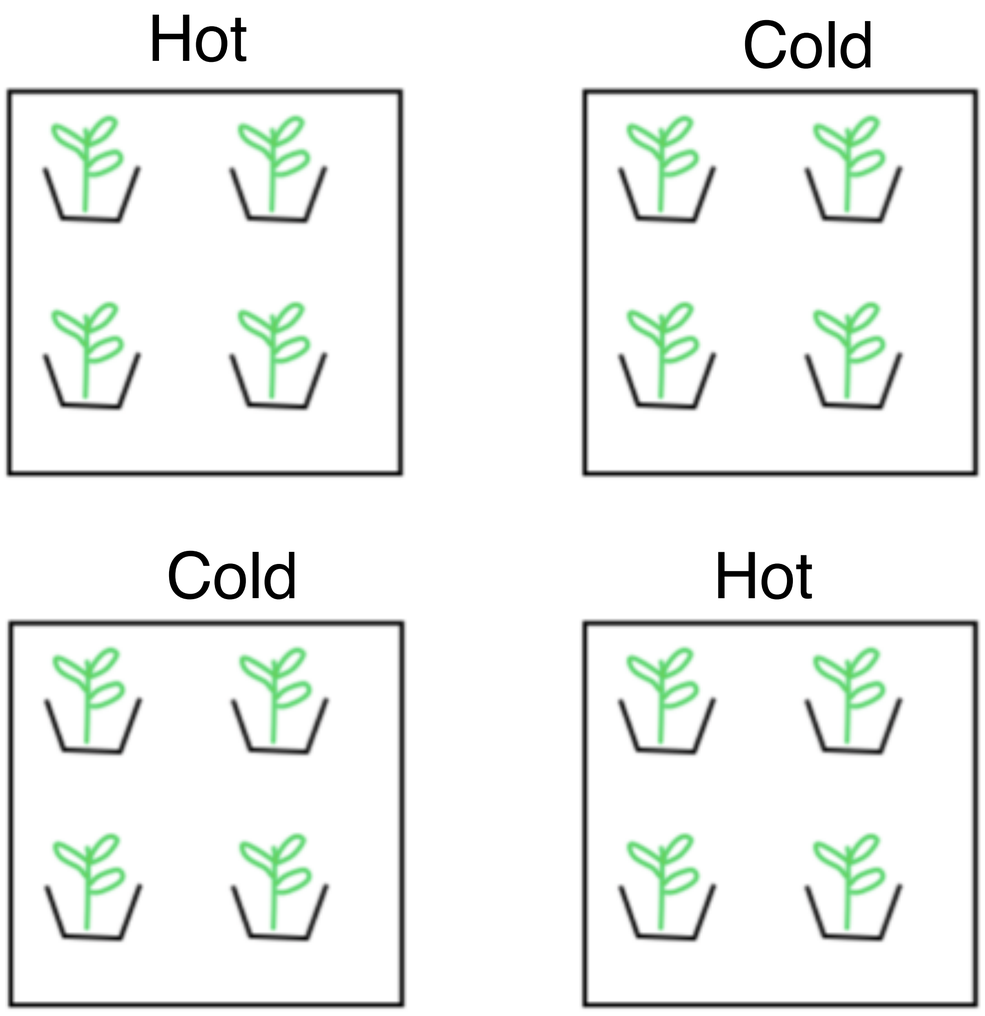


TABLE 1. Potential sources of confusion in an experiment and means for minimizing their effect.

Source of confusion	Features of an experimental design that reduce or eliminate confusion
1. Temporal change	Control treatments
2. Procedure effects	Control treatments
3. Experimenter bias	Randomized assignment of experimental units to treatments Randomization in conduct of other procedures "Blind" procedures*
4. Experimenter-generated variability (random error)	Replication of treatments
5. Initial or inherent variability among experimental units	Replication of treatments Interspersion of treatments Concomitant observations
6. Nondemonic intrusion†	Replication of treatments Interspersion of treatments
7. Demonic intrusion	Eternal vigilance, exorcism, human sacrifices, etc.

Hot, Cold measured at same time

Swap chambers?

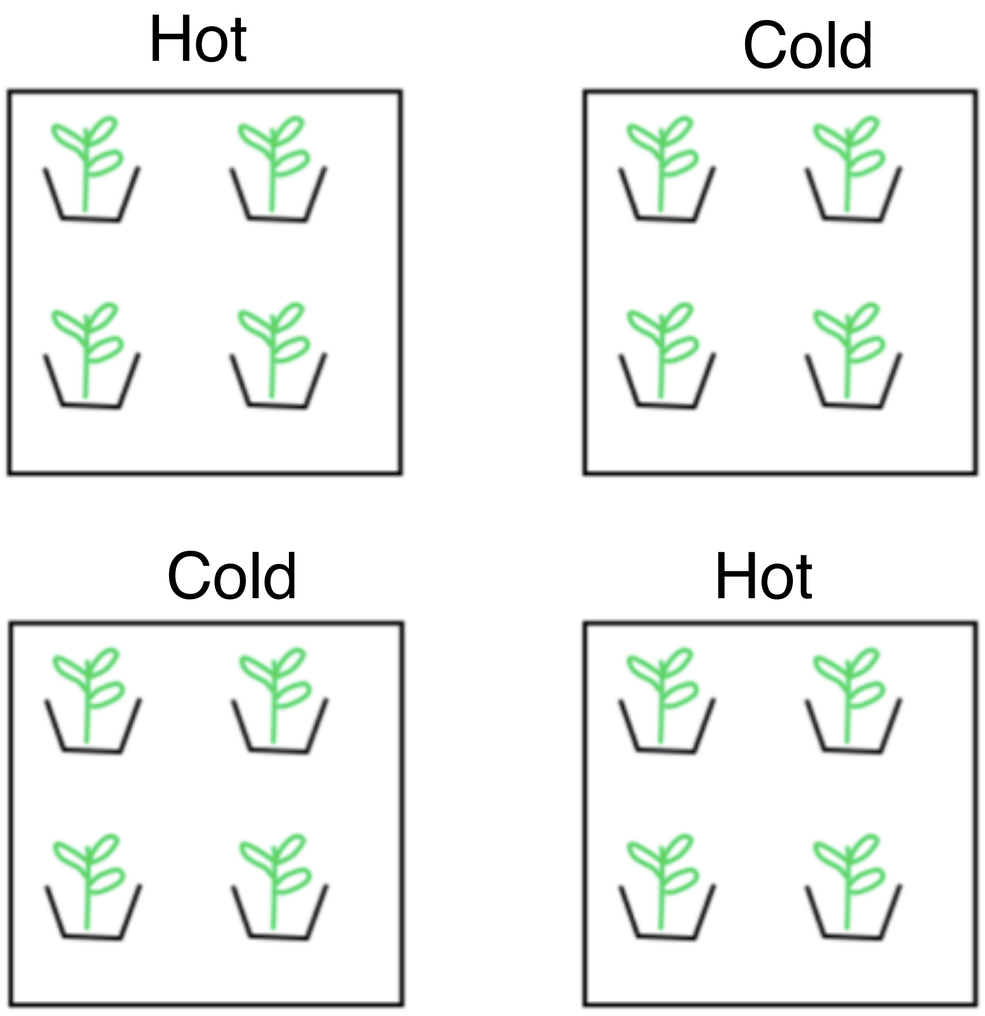
Randomization of chambers to temperatures

Replicate chambers

Chambers next to each other

Swap chambers

Confounding factors



Can we safely **interpret** the difference in expression between Hot and Cold as **caused** by temperature?

Is anything else consistently different in Hot vs Cold chambers?

Humidity?

Light intensity?

Water stress?

Growth rate?

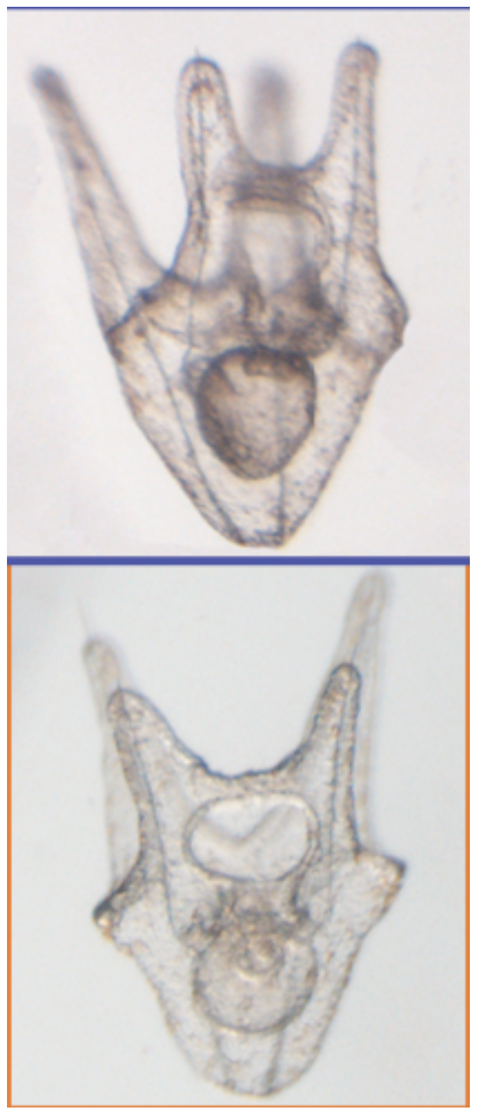
If you controlled these factors and found different results, would your interpretation change?

$\hat{\delta} = \hat{\mu}_{Hot} - \hat{\mu}_{Cold}$ estimates the **total** difference in expression between the two treatments.

This may be **more** than the **effect of temperature** *per se*

Example #2

To study the effect of Ocean Acidificiation on Sea Urchin larvae, researchers prepared 56 flasks with sea water and used independent CO_2 controllers to control the acidity of each flask



Are there problems with this experimental design?

Example #3

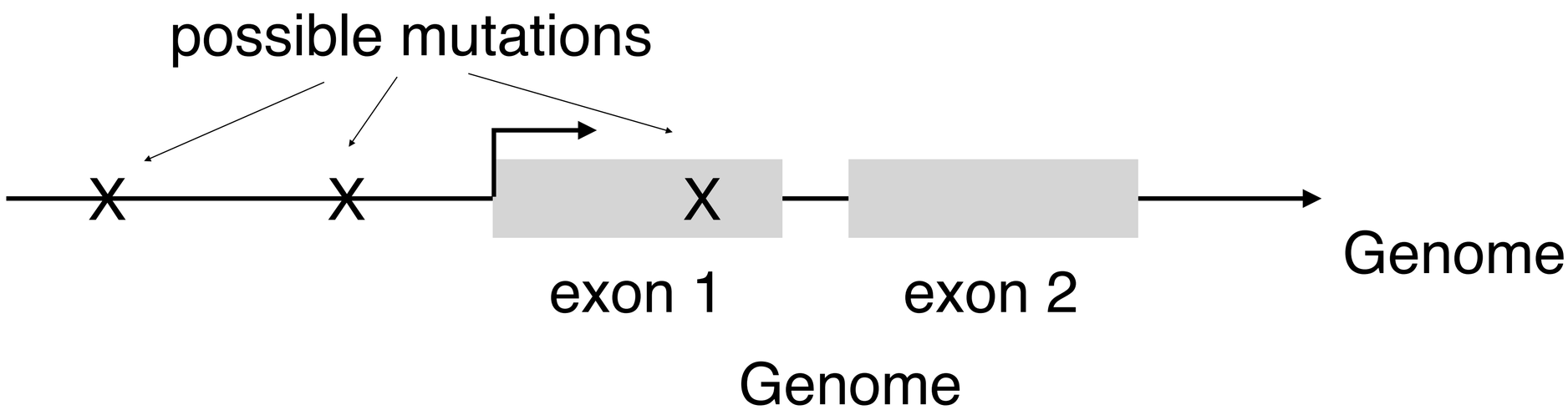
To study the effect of mutating the MC1R gene on fish fin colors, a researcher spends 2 years generating a knock-out mutant.

He places 6 fish of the wild-type strain in one tank and 6 fish of the mutant strain in a second tank.

When they get to 5cm in length, he measures the fin color of each fish

What is the experimental unit for the effect of MC1R on fin coloration?

He **TAGS 6 FISH** of the wild-type strain and 6 of the mutant strain and grows them **ALL IN THE SAME TANK**



Example #4

To study the timing of sugar signaling in *Arabidopsis* roots, roots were exposed to 100mM sucrose solution and the the concentration of starch was measured at 10 minute intervals

Assays were done in petri dishes with 5 seedlings each

10 petri dishes were started and a times 0 (before sucrose), 10min, 20min, 30min, and 40min, one randomly selected seedling was harvested from each plate and subjected to starch analysis.

What is the experimental unit for the effect of sucrose at 10min? 20min?

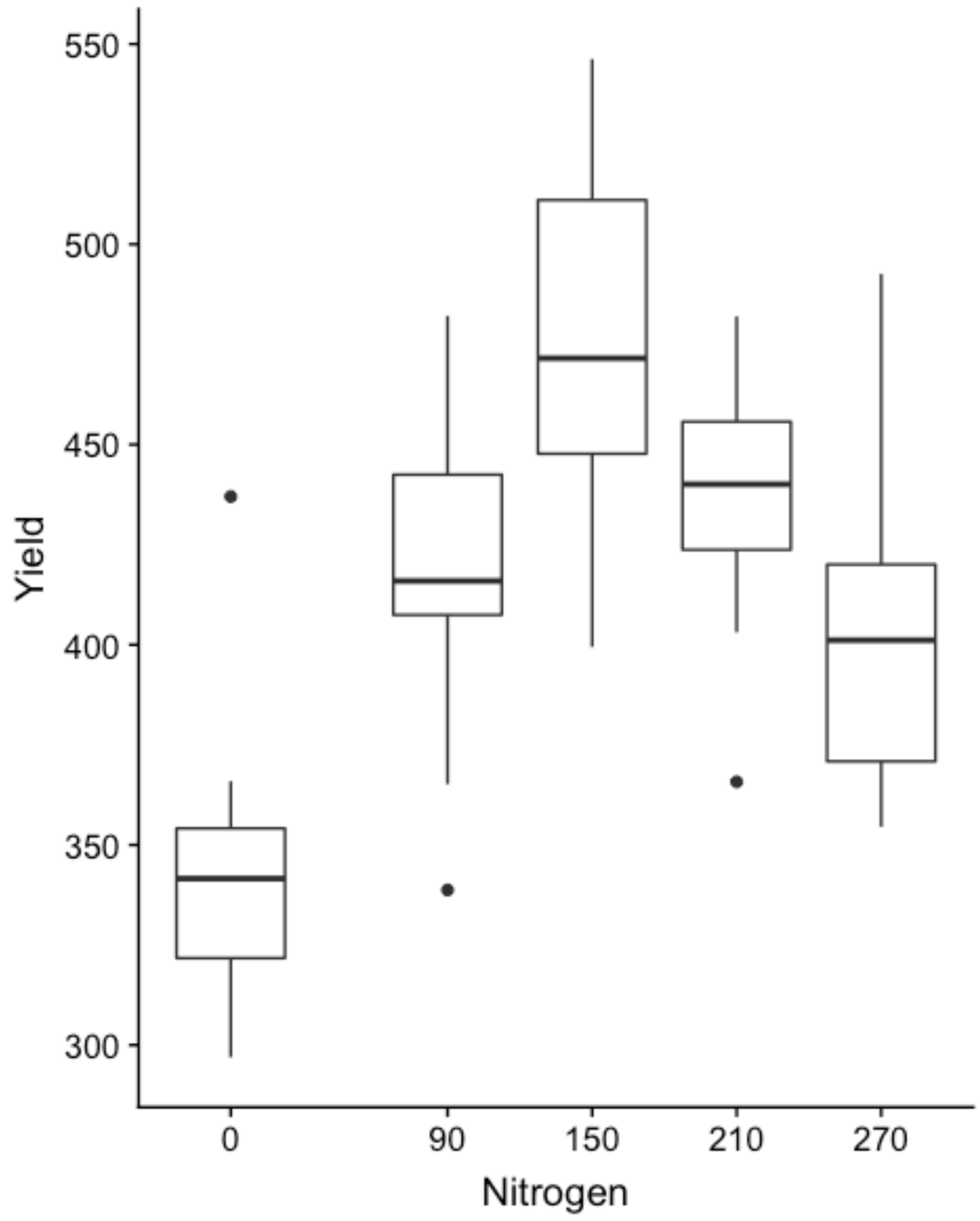
Example #5

An experiment was run to evaluate effects
of increased nitrogen fertilization
on tuber yield of frying potatoes

5 nitrogen regimes (applied to plots):
0, 90, 150, 210, 270 lbs / acre at emergence

10 reps / treatment combination

Response: total yield per plot



How were the +N treatments applied?

What is needed to ensure that the 10 reps / level are **independent**?