Analyze the Pulse experiment

What Research Question should we target?

What analysis should we do?

In what way might we argue this is a **mensurative experiment**?

Analyze the Pulse experiment

What Research Question should we target?

What analysis should we do?

In what way might we argue this is a **mensurative experiment**?

Scope: All UCD grad students at any time

~mensurative (no generalization)

We’ve really only observed an effect 1x

Scope: All UCD grad students **last Tuesday at 10:30am**

manipulative

If we observe an effect with high confidence we can state confidently that standing has an effect **at least in some contexts**

```
data:  standing_obs and sitting_obs
t = 0.49948, df = 64, p-value = 0.6192
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -5.674927  9.458711
sample estimates:
mean of x mean of y
 79.00000  77.10811
```

What do these numbers mean?

What is a Treatment Effect?

How much did standing affect pulses **on average** in the context of Class 1?

Observations:

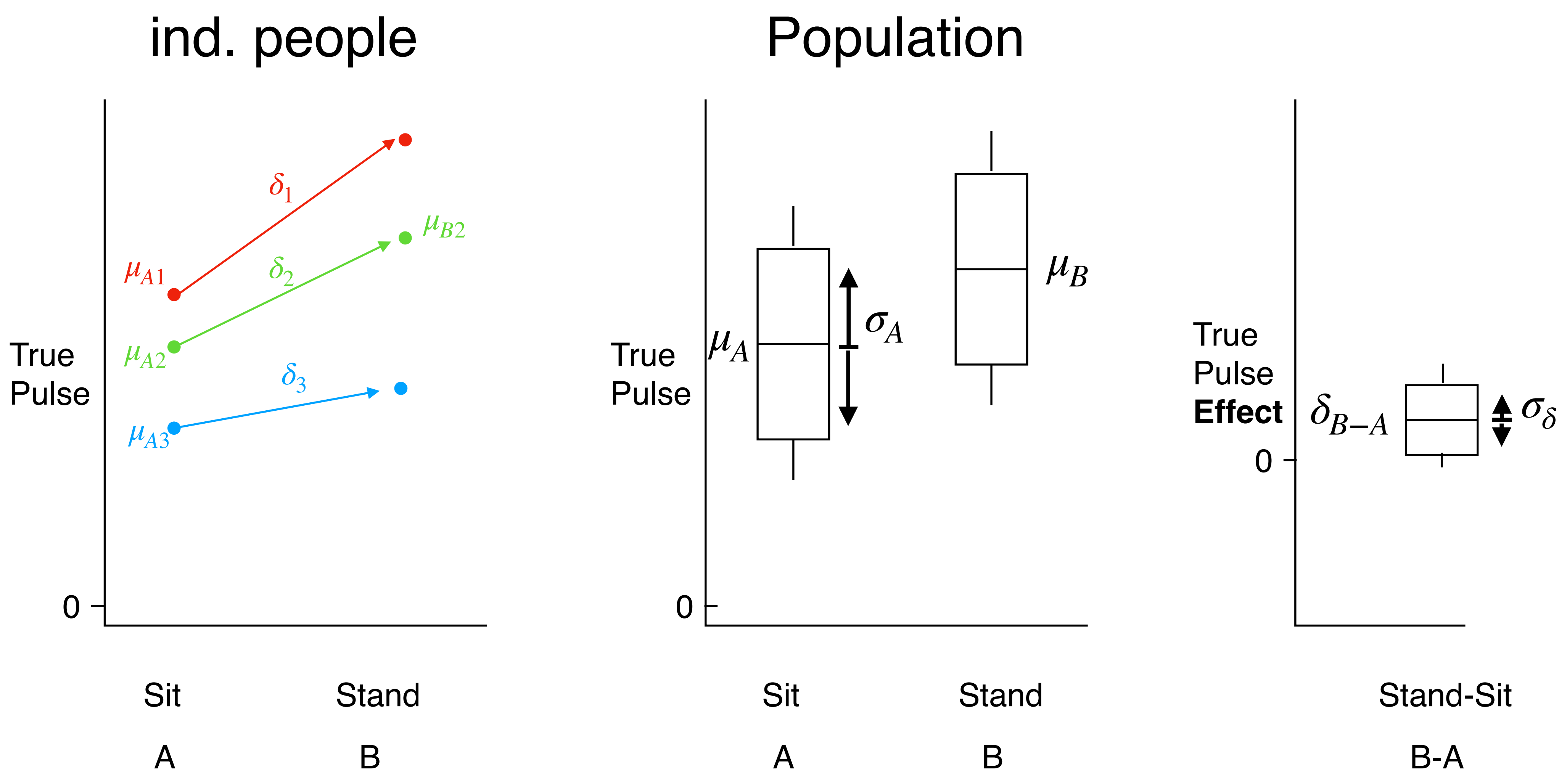
Everyone started with a different pulse while sitting

The change when standing was different for each person that stood

Each observation had measurement error

It’s not useful (scientifically) to report the effect for each person

What is a Treatment Effect?



Average (mean) units: bpm

$$\mu_A = \frac{\mu_{A1} + \mu_{A2} + \mu_{A3} + \ldots \mu_{AN}}{N} = \frac{\sum \mu_{Aj}}{N}$$

Variance (Average deviation^2)

$$\sigma_A^2 = \frac{\sum (\mu_{Aj} - \mu_A)^2}{N}$$

units: bpm^2

$$\sigma_\delta^2 = \frac{\sum (\delta_j - \delta)^2}{N}$$

SD (Standard deviation) $\sigma_\delta = \sqrt{\sigma_\delta^2}$

~ average abs(deviation)

units: bpm

μ_{A1} Pulse of person **1** in treatment_level **A**

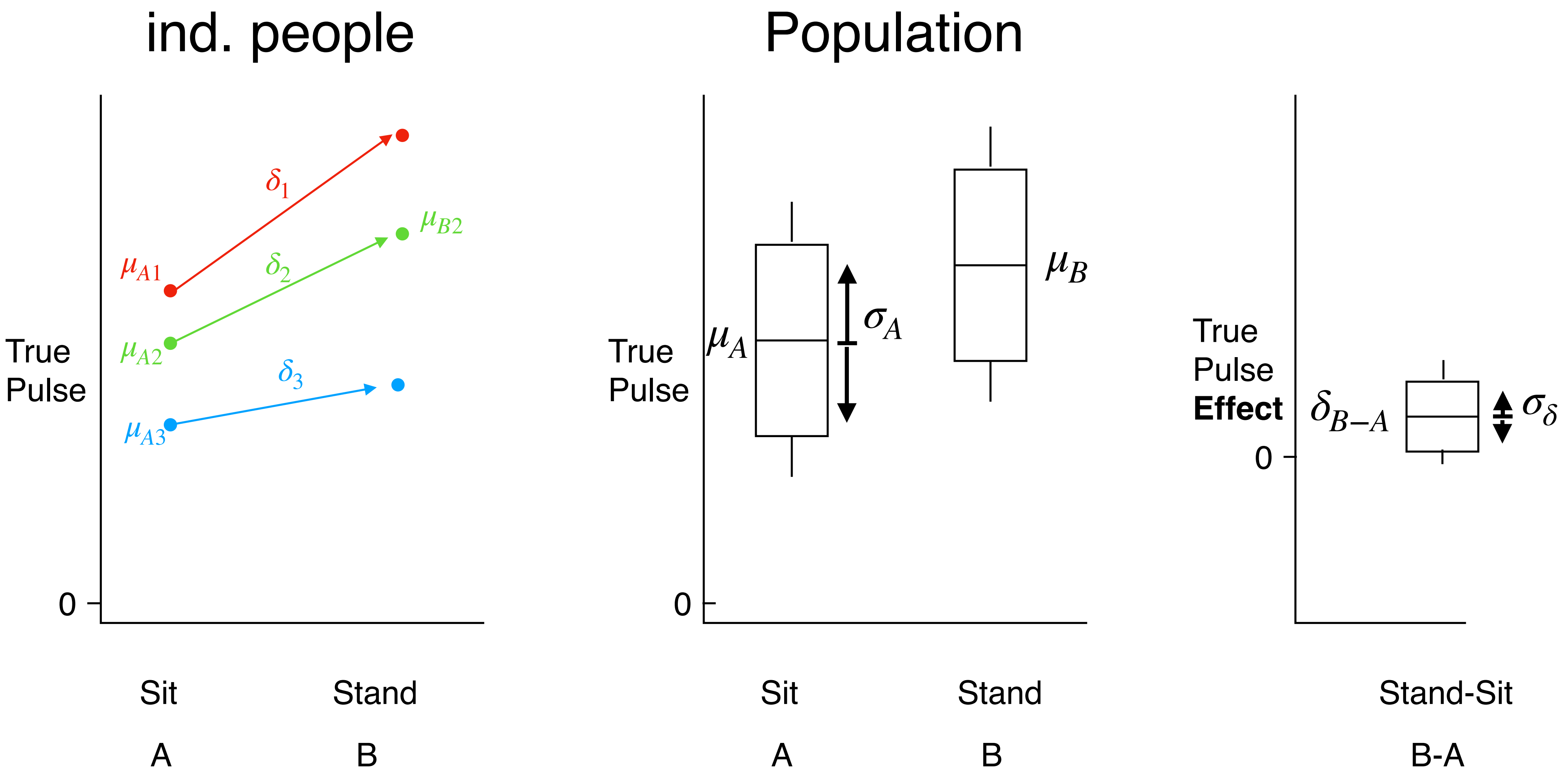
δ_1 **Effect of standing** for person **1**

μ_A **Average** pulse in treatment_level **A**

δ_{B-A} **Average** effect of standing

These are all **TRUE values** They exist, but we can't know them

Why focus on the Population?

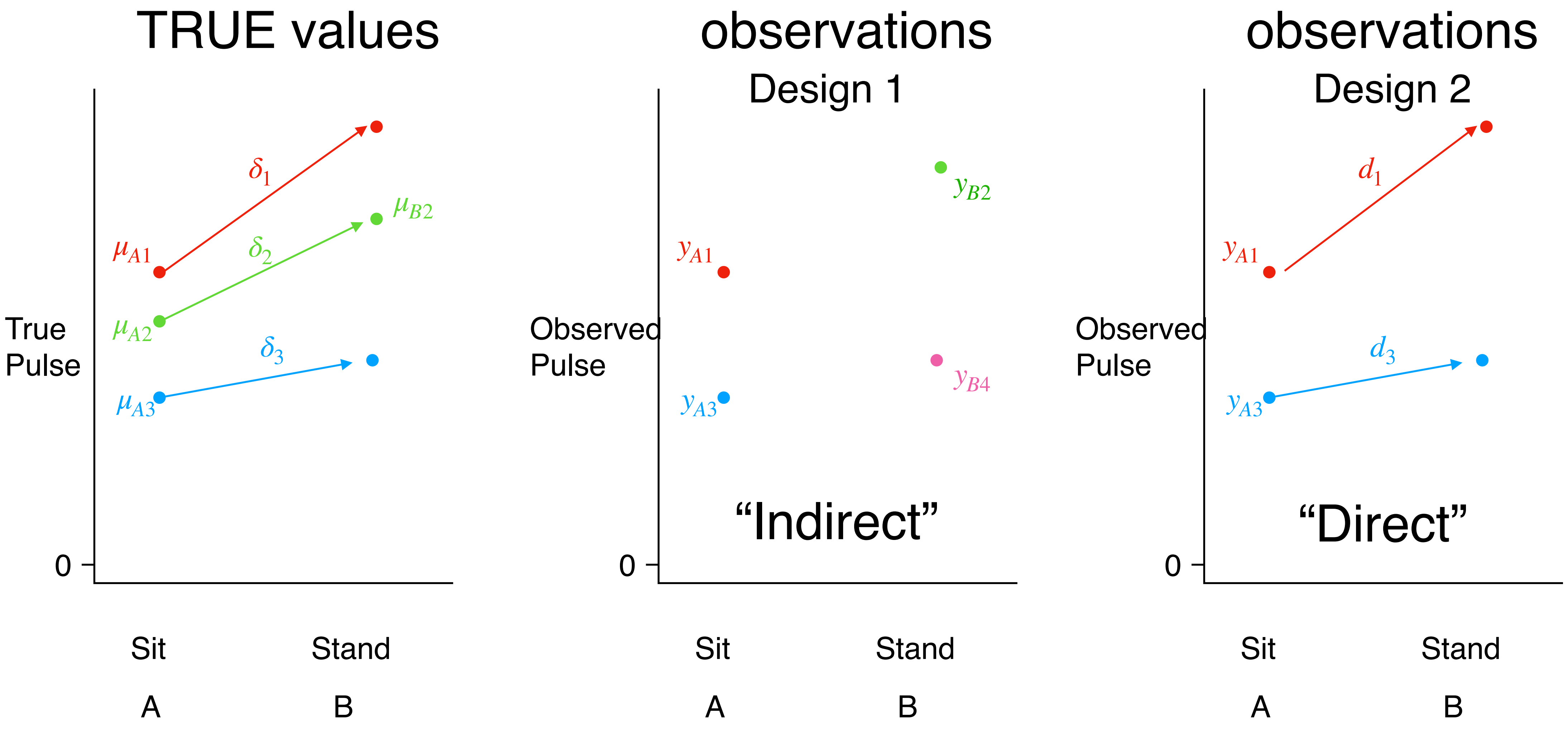


If we learn the **population mean and variance** …

We can generalize to **new individuals**

"our" individuals are just examples to learn from

We can't observe the population



y_{A1} **Observed pulse** of person **1** in treatment_level **A**

d_1 **Observed effect of standing** for person **1**

We only observe **some of the population**

Each observation has error $y_{A1} - \mu_{A1} = \epsilon_{A1}$

We use the observations to estimate the TRUE values

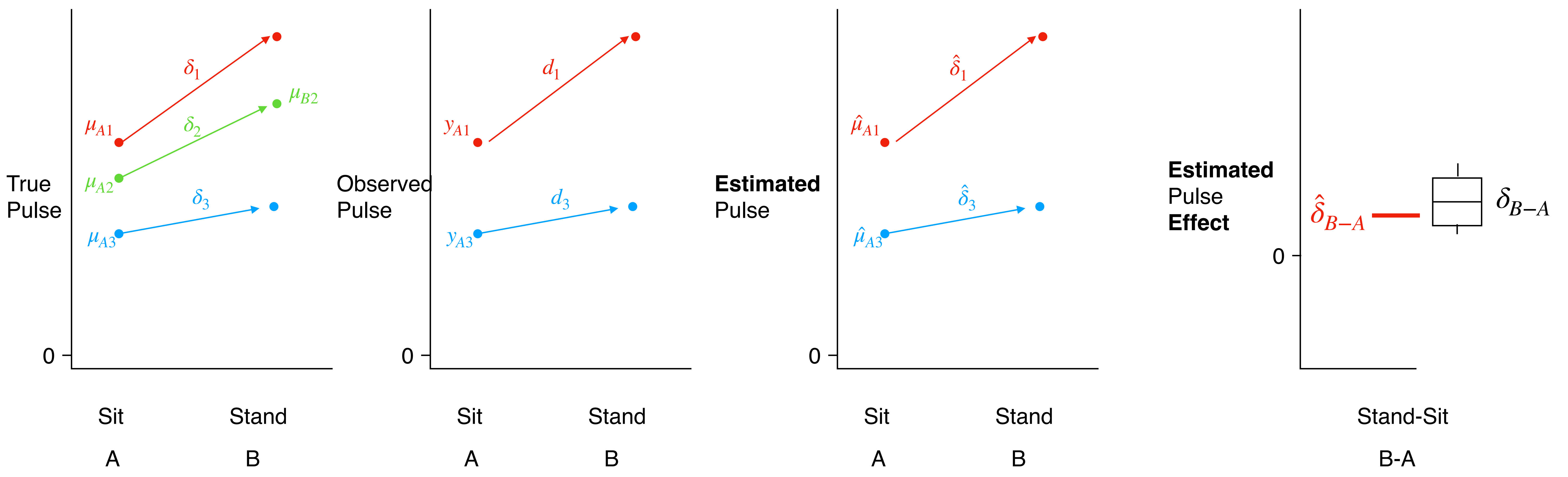
ind. values

pop. values

TRUE values

observations

observations



$\hat{\mu}_{A1}$ **Estimated** Pulse of person **1** in treatment_level **A**

$\hat{\delta}_1$ **Estimated effect of standing** for person **1**

$\hat{\mu}_{A1} = y_{A1}$ $\hat{\delta}_1 = d_1$

$\hat{\delta}_{B-A}$ **Estimated** average effect of standing

$$\hat{\delta} = \frac{\sum \hat{\delta}_j}{n}$$

This procedure inherently has error

Our estimate *will not* equal the true value: $\hat{\delta}_{B-A} - \delta_{B-A} = \epsilon$ We call ϵ our **error**

Causes of error:

Sampling error: we didn't sample the whole population

Measurement error: we each observation was noisy

We can't know ϵ. But we can state the **typical size** of ϵ^2 *for **Direct estimates***

$$\frac{\text{Variance of population + Variance of measurements}}{\text{Sample size}} = \frac{\sigma_\delta^2 + \sigma_m^2}{n} = \sigma_r^2(\hat{\delta})$$

$\sigma_r(\hat{\delta})$ is called **Standard Error**

~ **Average error** if you were to repeat the experiment many times

Goal: Create experiments that minimize the standard error

Summary

Notation:

Greek letters ($\mu, \delta, \sigma, \epsilon$) represent TRUE values

Roman letters (y, d, e, s, n) represent observation, data

Greek letters **with hats** ($\hat{\mu}, \hat{\delta}$) are estimates (calculated from data to estimate TRUE values)

Subscripts represent subgroups: (letters are treatments, numbers are replicates)

Concepts:

Goal: Learn about population means

Each experimental result has an associated error

The average size of this error is a property of the experimental design (standard error: $\sigma_r(\hat{\delta})$)

*for **Direct estimates***

$$\sqrt{\frac{\text{Variance of population + Variance of measurements}}{\text{Sample size}}}$$

Prepare for Tuesday:

Read Hurlbert 1984 up to “Interspersion of Treatments”

Questions:

Was the Pulse Experiment **sufficiently controlled** so that we can interpret our treatment effect estimate as a valid estimate of **the effect of standing**?

There are 3 ways to reduce the standard error in this experiment.

$$\sigma_r(\hat{\delta}) = \sqrt{\frac{\text{Variance of population} + \text{Variance of measurements}}{\text{Sample size}}}$$

Should we try to reduce all 3?