Experimental Design

What is an experiment?

Understand two things:

Purpose & Process

Hugely Important Ideas

Purpose of an experiment:

Establish causation.

(Purpose of an observational study:

Establish a relationship.)

A really simple process looks like this:

1) You think something, call it **A**, causes something else call it **B**.

2) So, you CAREFULLY DESIGN A PROCEDURE where

you do **A** to **certain victims** AND **A** is the ONLY thing that happens to **those victims**

AND you pick **other victims** who are just like the **certain victims**

you **DON'T** do **A** to **other victims**.

3) Then you can say,

"Anything that happened to the **certain** victims happened because of **A.**"

{You secretly hope the result is **B**}

Correct language.

The key word:

VARIABLE

It get's used in weird ways in discussing experiments.

We think as variables as measurement.
In an experiment variables are:

- Any influence on or thing that affects the subject
  (That you do on purpose or not)
- Are often binary

ANY INFLUENCE

The brightness of the light in the room.
The gender of the participant.
The food they had for breakfast

OFTEN BINARY

Being called an "asymptote" OR not is a variable.
Received medication: yes or no?
Treated: Yes or no - often the most important variable.

A good experiments
- "Controls" all variables:

- Make all variables the same for every subject, except the treatment.

<u>Control these variable to be the same for ALL subjects</u>

- The temperature in the room.
- The time of day.
- The age of the subject.
- The day of the week.
- The body odor of the person applying the treatment.

....

<u>Make one variable different on purpose:</u>

- Different subjects get different specific treatments.

<u>What you can conclude</u>?

Any different results must be due to difference in treatments.

A good experiments

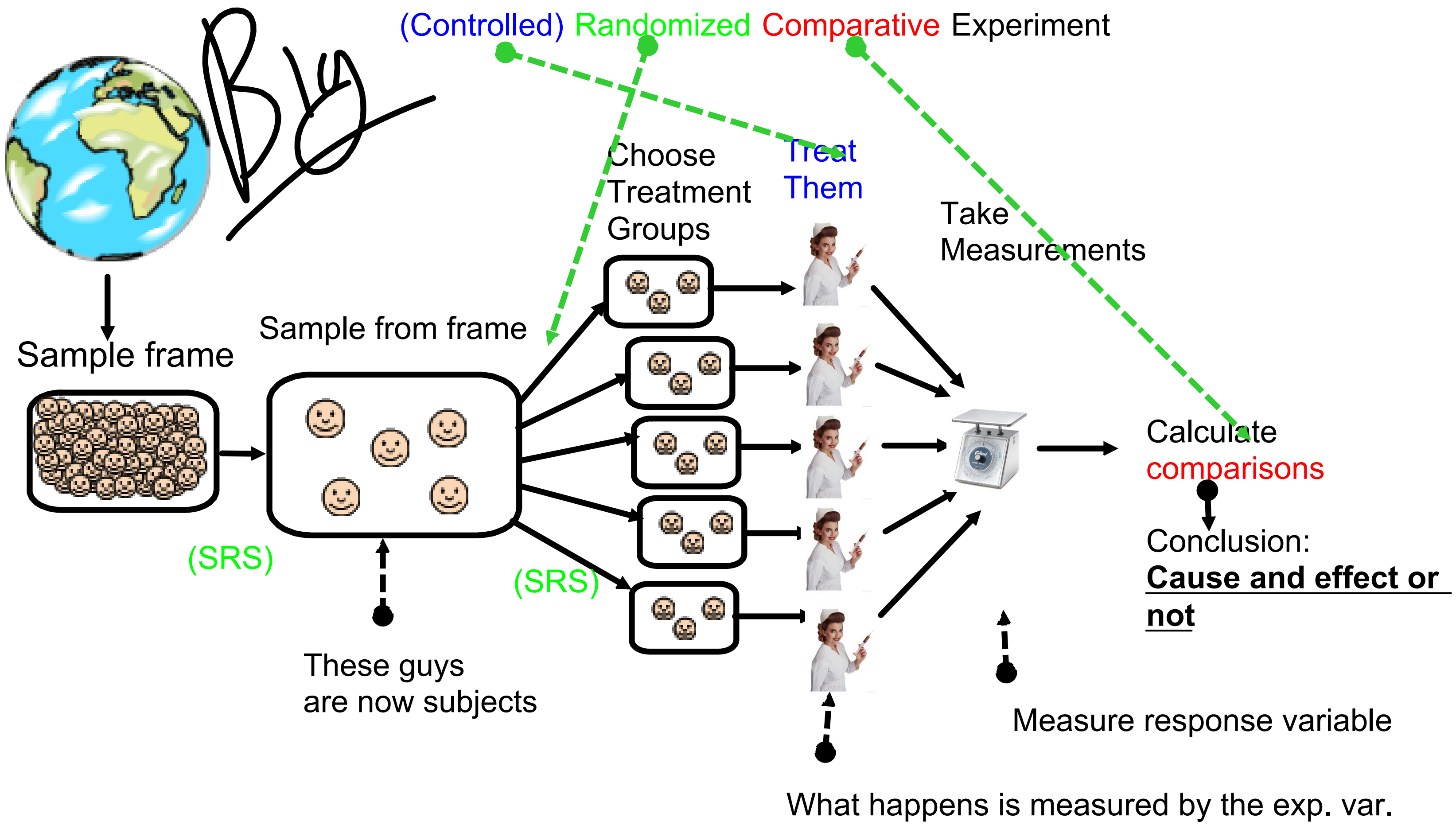
- "Controls" all variables:

- Make all variables the same for every subject, except the treatment.

You hope it explains what happens to the response variable

Types of variables:

The result you measure at the end is the **response variable**.
The variable, or treatment, you are testing is the **explanatory variable**.
The **treatment** is any action or condition you apply to the subjects.
The **subjects** are who or what you treat.



(Controlled) Randomized Comparative Experiment
Big
Sample frame
Sample from frame
(SRS)
(SRS)
These guys
are now subjects
Choose
Treatment
Groups
Treat
Them
Take
Measurements
Calculate
comparisons
Conclusion:
Cause and effect or not
Measure response variable
What happens is measured by the exp. var.

(Controlled)
- Control variables
- Have a control group

Randomized
- Random selection of subjects. if possible.
- Random assignment to treatment groups.

Comparative
- Comparing response variable across groups

Replication: Make N as big as possible

Logic

How do you control variables? Options:
- Actually control them: Make the circumstance of the variable the same for all treatment groups.
- Isolate the effect: Make treatment groups that are different for that variable.
- Leave it to luck: Make your groups so large difference along lurking variables will average out.

Why do you randomize?
- Remove human influence and potential bias.
(Keep humans including you, from making choices of any kinds)

What does replication do? (the book ignores part of this):
- Book discusses: Lots of subjects average out random variation.
- Book does not discuss: Do the experiment again.

So,

With control of variables,
protecting from biases with randomization,
comparing to a control group (and may be others)
using a large enough set of subjects to average out lurking variables,
we can conclude the different results are due to the difference in treatment.

(We conclude causation..

Problems:Controlling **<u>all</u>** the variables is hard.

- You have to think of/identify ALL of them
- You have to be able to do something about ALL of them

Names for problems:

Lurking variables
Confounded variables

Lurking variables - YOUR FAULT

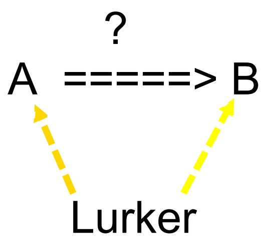
You missed it in the list of variables you managed.

1) Either you didn't think of it.

2) You thought of it but didn't* measure/monitor it.

What's it do? If you didn't manage it, you can't say it didn't cause the result.

After the study: You can't confidently say your treatment caused the result.



A happens, B happens. A result!

But maybe:

Lurker happened so A happened

&

Lurker happened so B happened

* Chose not too, didn't have the resources

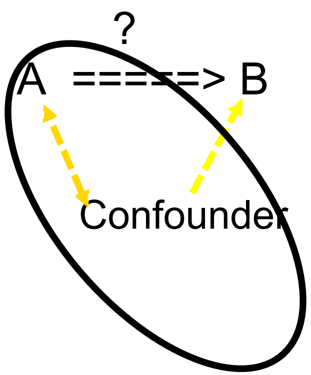
Fire, teacher, seating

Confounding variables -NOT YOUR FAULT

The confounded variable can not be seperated.

What's it do? If you didn't manage it, you can't say it didn't cause the result.

After the study: You can't confidently say your treatment caused the result.



A happens, B happens. A result!

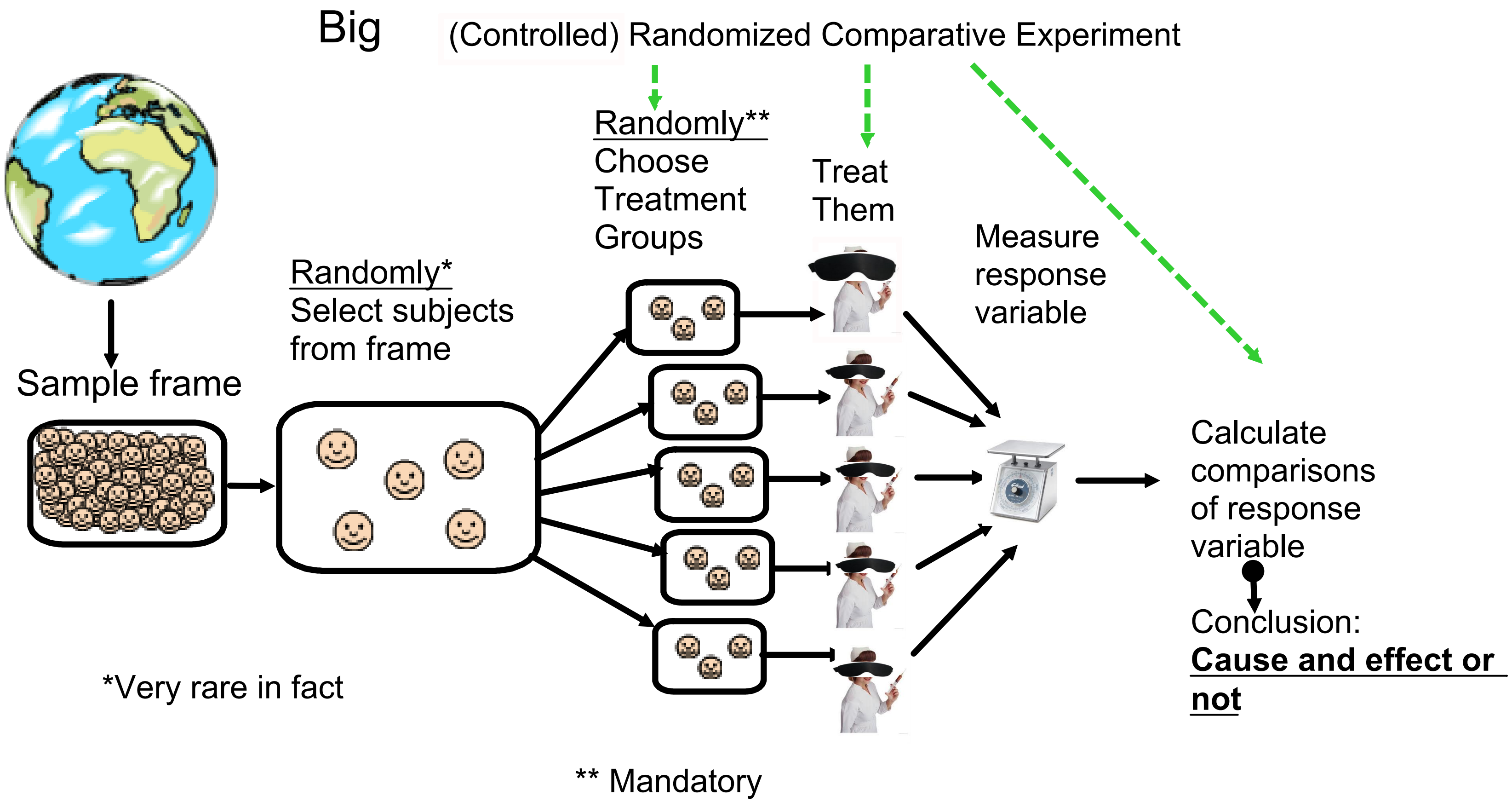
Maybe confounder happened so B happened.

* Chose not too, didn't have the resources

Is it a big randomized controlled comparative experiment?

Explanatory variable response variable.

Other variables discussed.



Big
(Controlled) Randomized Comparative Experiment
Randomly** Choose Treatment Groups
Treat Them
Measure response variable
Sample frame
Randomly* Select subjects from frame
Calculate comparisons of response variable
Conclusion:
Cause and effect or not
*Very rare in fact
** Mandatory

Random and blind

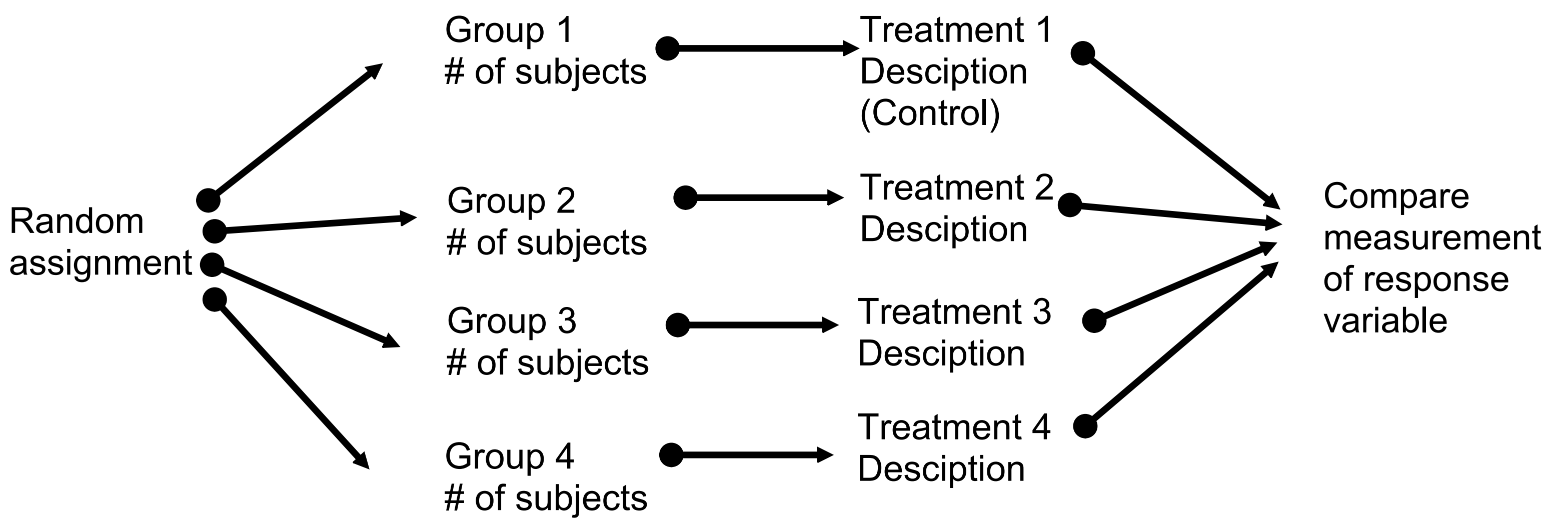
"Blind study"

Double blind, think triple blind

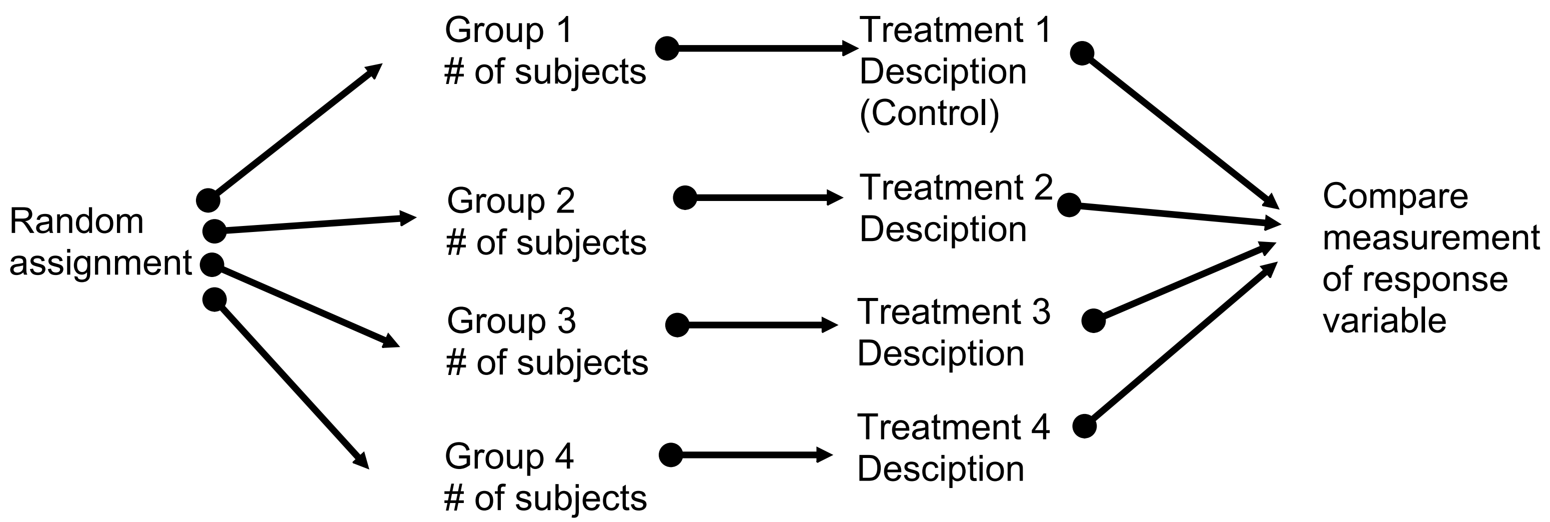
Double blind: Neither the subject or the experimenters know which subject is getting which treatment.

"Triple blind" The subject, the person applying the treatment, and the person measuring the response variable do not know which subject is getting which treatment.

How to make a diagram



How to make a diagram



(Controlled)
- Control variables
- Have a control group

Randomized
- Random selection of subjects. if possible.
- Random assignment to treatment groups.

Comparative
- Comparing response variable across groups

Logic

How do you control variables? Options:
- Actually control them: Make the circumstance of the variable the same for all treatment groups.
- Isolate the effect: Make treatment groups that are different for that variable.
- Leave it to luck: Make your groups so large difference along lurking variables will average out.

Why do you randomize?
- Remove human influence and potential bias.

(Keep humans including you, from making choices of any kinds)

What does replication do? (the book ignores part of this):
- Book discusses: Lots of subjects average out random variation.
- Book does not discuss: Do the experiment again.

So,

With control of variables,
protecting from biases with randomization,
comparing to a control group (and may be others)
using a large enough set of subjects to average out lurking variables,
we can conclude the different results are due to the difference in treatment.

(We conclude causation..

But....

By twos, pick one and design and experiment.

We want to study if taking Airborne prevents colds.
Design an experiment. You have 320 subjects.

We want to study if thanking your teacher everyday improves your grade.
Design an experiment. You have 320 subjects.

We want to study if the condition of the top of a HS guy's head (hair style vs. wearing a hat all day vs. ?) effects dating. You have 320 subjects.

Skills/knowledge so far:

Know what makes an experiment:
- Large: replication or big n
- Controlled: Controlling all variables
  (Large n, actual control, blocking)
- Random
- Comparative

Vocab
- Explanatory variable
- Response variable
- Individual
- Subjects
- Treatments

Drawing a complete basic diagram

How big a difference is big enough?

Average writing section SAT scores:
Ludlowe: 502
Warde: 500
Would you say Ludlowe is making students who are better at the SAT?
If we took the tests again, or next year, would you have confidence Ludlowe score will be higher?

510 vs 500? 520 vs 500? 550 vs 500?

How big is big enough?

510 vs 500? 520 vs 500? 550 vs 500?

How big is big enough?

Big enough to be statistically significant.

How big is that? Your choice.

Calculate a 95% confidence interval, outside the interval is too big / a real difference.

Calculate a 95% confidence interval, outside the interval is too big / a real difference.
But this is not a proportion - different formula.

Say these are the scores:

Warde: 500 Ludlowe: 525

And this is the 95% confidence interval for the Warde score:

{487.71, 512.29}

Significantly different?

Yes: Getting answer this far away has less than a 5% chance.

But if this: {472.75, 537.25}

Significantly different?

No: In the 95% range, it is not different enough to say the parameters for the groups are different.

A difference in two statistics for two groups is ***statistically significant***
if the difference so large
that statistical calculations tell us
a difference that large
would not happen very often
if the groups were the same.

A difference in two statistics for two groups is not ***statistically significant***
if the difference so small
that statistical calculations tell us
a difference that small
could happen often
even when the groups are really the same.

Explain what these statements mean:

The difference in math SAT scores between Central and Staples was 10 points. This was significantly different. What does that mean?

The difference in math SAT scores between Central and Staples was 10 points. This was not significantly different. What does that mean?

The difference in number of colds between Airborne users and non-Airborne users was not statistically significant.

The difference the response variable is so {small/large} that statistical calculations tell us a difference that {small/large} is {quite likely/quite unlikely} if the parameters are really same.

Blocking:

When you can identify, in advance, a variable or feature of the subjects that you think will make one set of subjects respond differently.

Gender
Age
Experience
Interests
Physical traits
Culture
Species

It allows you to draw a separate conclusoin for each group/block.

Then group people who are alike and treat them

and THEN pick
treatment groups by SRS

I invented a new biceps building exercise. I randomly picked 20 subjects. One group did the exercise, one group did the best known exercise. I measured the gain in biceps mass.

Treatment group		Control group	
	Gain:		Gain:
1	2 m	11	1 m
2	4 m	12	2 m
3	1 m	13	3 m
4	3 m	14	2 m
5	6 m	15	4 m
6	3 m	16	3 m
7	8 m	17	2 m
8	4 m	18	4 m
9	1 m	19	2 m
10	5 m	20	1 m
Average gain: 3.7 m		Average gain: 2.4 m	

I do stat. math and find this is not a significant difference. But...

I invented a new biceps building exercise. I randomly picked 20 subjects. One group did the exercise, one group did the best known exercise. I measured the gain in biceps mass.

Treatment group men:

Gain:

1	2 m
2	3 m
3	1 m
4	3 m
5	1 m

Average gain: 2 m

Treatment group women:

6	4 m
7	8 m
8	4 m
9	6 m
10	5 m

Average gain: 5.4 m

Control group men:

Gain:

11	1 m
12	2 m
13	3 m
14	2 m
15	4 m

Average gain: 2.4 m

Control group women:

Gain:

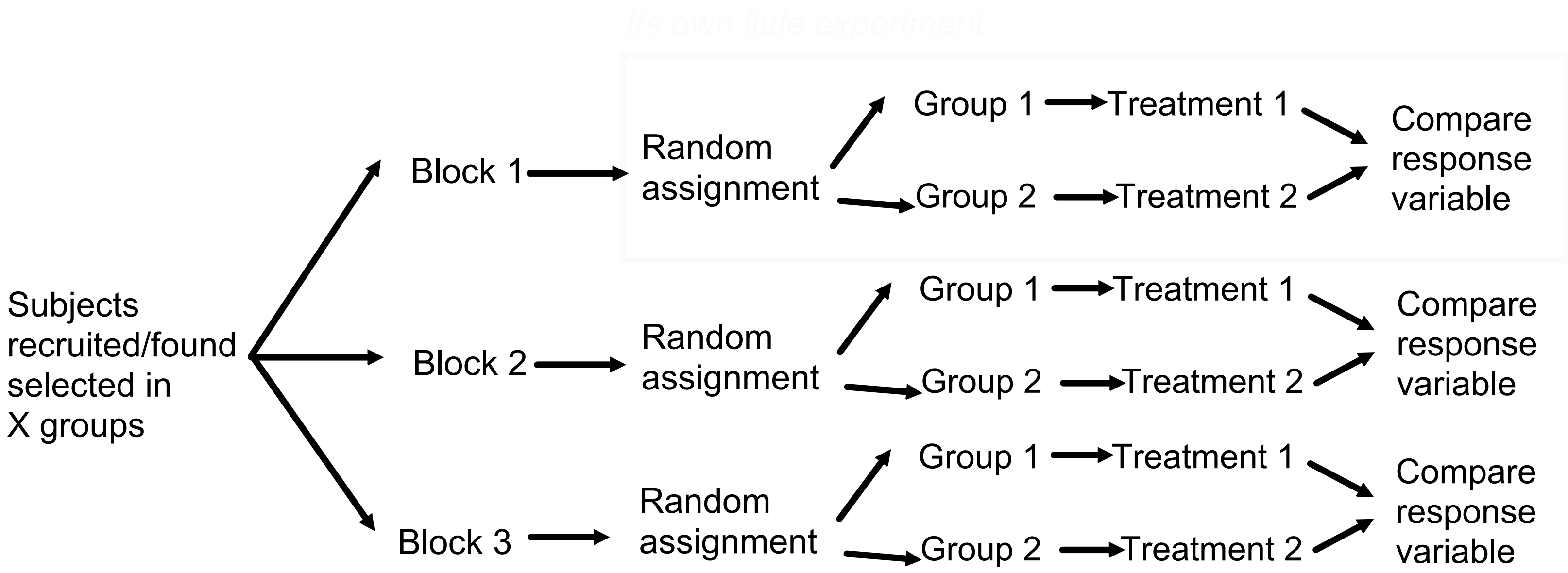
16	3 m
17	2 m
18	4 m
19	2 m
20	1 m

Average gain: 2.4 m

I do stat. math and find men are not significant, and was worse.

Women were significant.

Basically you do the same experiment
but on different groups

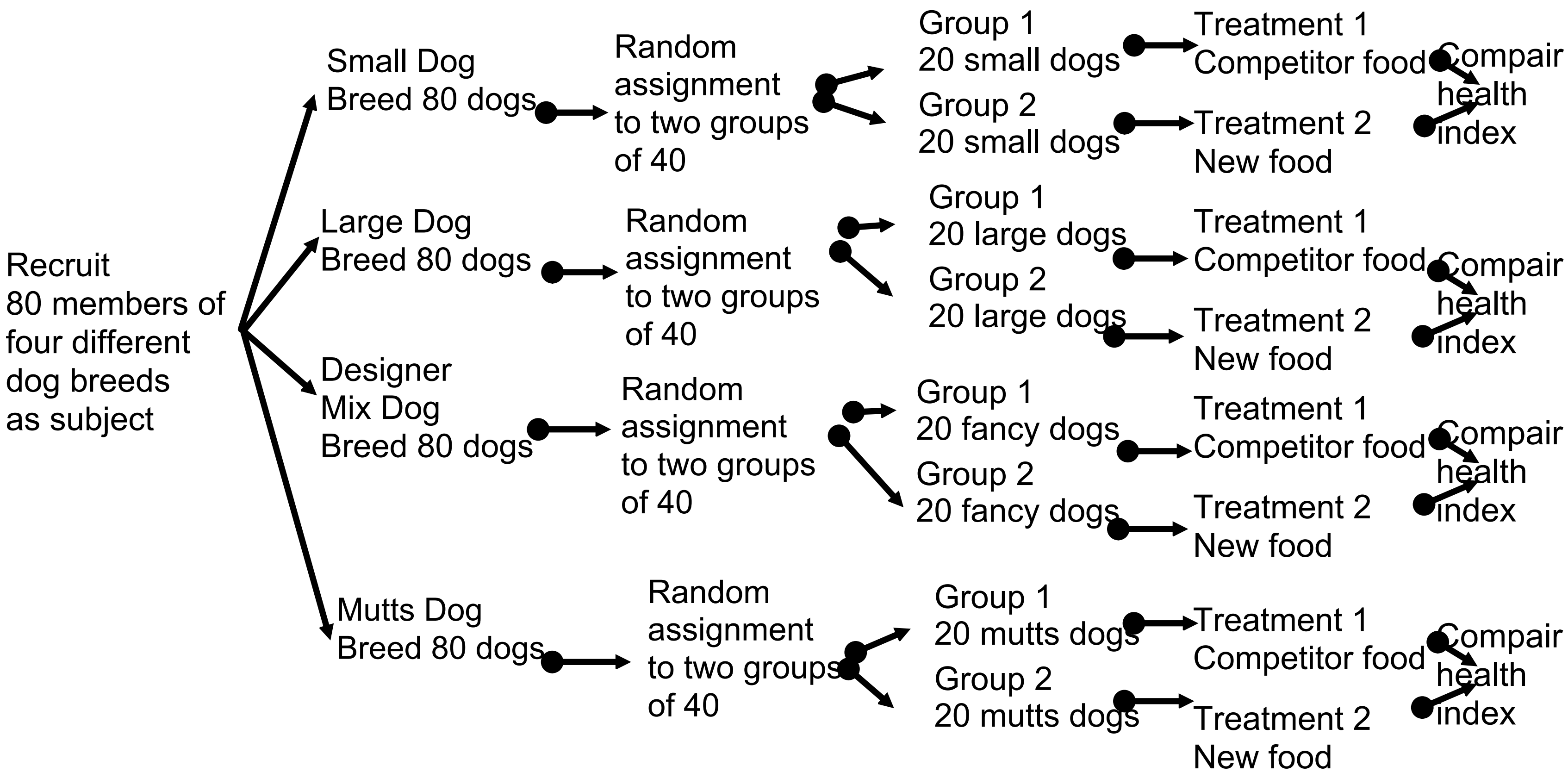

Its own little experiment
Subjects recruited/found selected in X groups
Block 1
Block 2
Block 3
Random assignment
Random assignment
Random assignment
Group 1
Group 2
Treatment 1
Treatment 2
Compare response variable
Group 1
Group 2
Treatment 1
Treatment 2
Compare response variable
Group 1
Group 2
Treatment 1
Treatment 2
Compare response variable

Nutritional value of a dog food.

Marketing research on how successful a celebrity magazine would be.

People's willingness to change to a new video format.

Nutritional value of a dog food. You have a budget for 360 subjects



Skills:

Explain the difference between an observational study and an experiment

Identify lurking or confounded variables.

Identify explanatory and response variables

Explain what statistical significance in the context of a problem.

Diagram 2 kinds of studies

- Regular Controlled Random Experiment
- Block

Decide when blocking

Design experiments with or without blocking

Identify what makes a study

- Random
- Controlled
- Comparative
- Double blind

Criticize a study for not being

- Random
- Controlled
- Comparative
- Double blind (3 ways)