

ASSIGNING PARTICIPANTS TO CONDITIONS

Scott Klemmer

Should every participant use
every alternative?

Which vacuum cleaner is more effective?



**Draw vacuum
cleaner**

What are the measures?

What are the measures?

- Faster?

What are the measures?

- Faster?
- Cleaner?

What are the measures?

- Faster?
- Cleaner?
- Fatigue?

What are the measures?

- Faster?
- Cleaner?
- Fatigue?
- Ergonomics?

What are the measures?

- Faster?
- Cleaner?
- Fatigue?
- Ergonomics?
- Environmental impact?

What are the measures?

- Faster?
- Cleaner?
- Fatigue?
- Ergonomics?
- Environmental impact?
- Portability?

What are the measures?

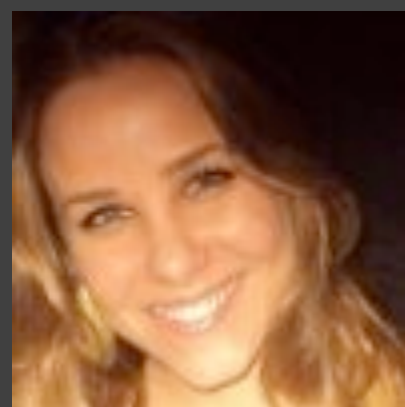
- Faster?
- Cleaner?
- Fatigue?
- Ergonomics?
- Environmental impact?
- Portability?
- Ease of use?

What are the measures?

- Faster?
- Cleaner?
- Fatigue?
- Ergonomics?
- Environmental impact?
- Portability?
- Ease of use?
- Fewer errors?

- Manipulation: Vacuum type

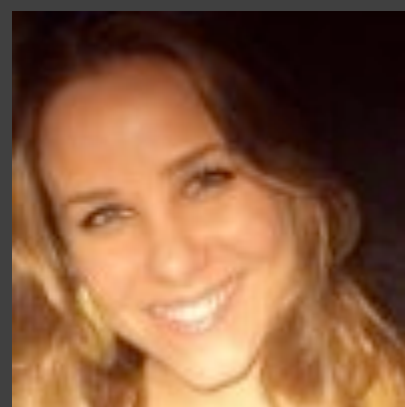
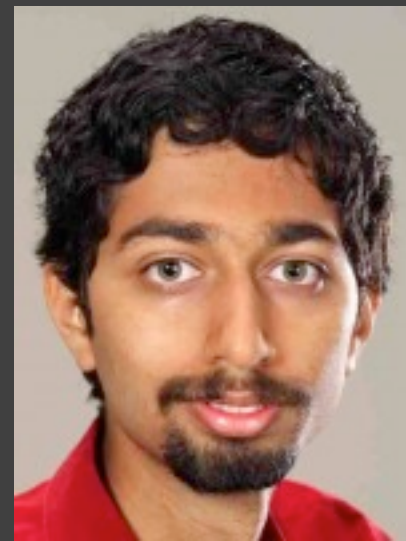
- Manipulation: Vacuum type
- Measure: Speed, cleanliness



Between subjects design

Half the participants use one interface

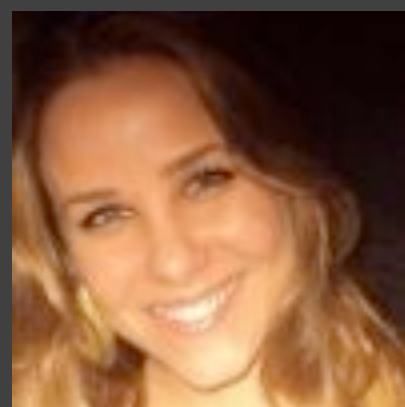
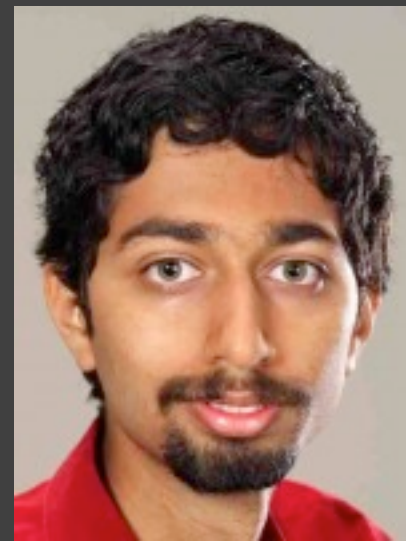
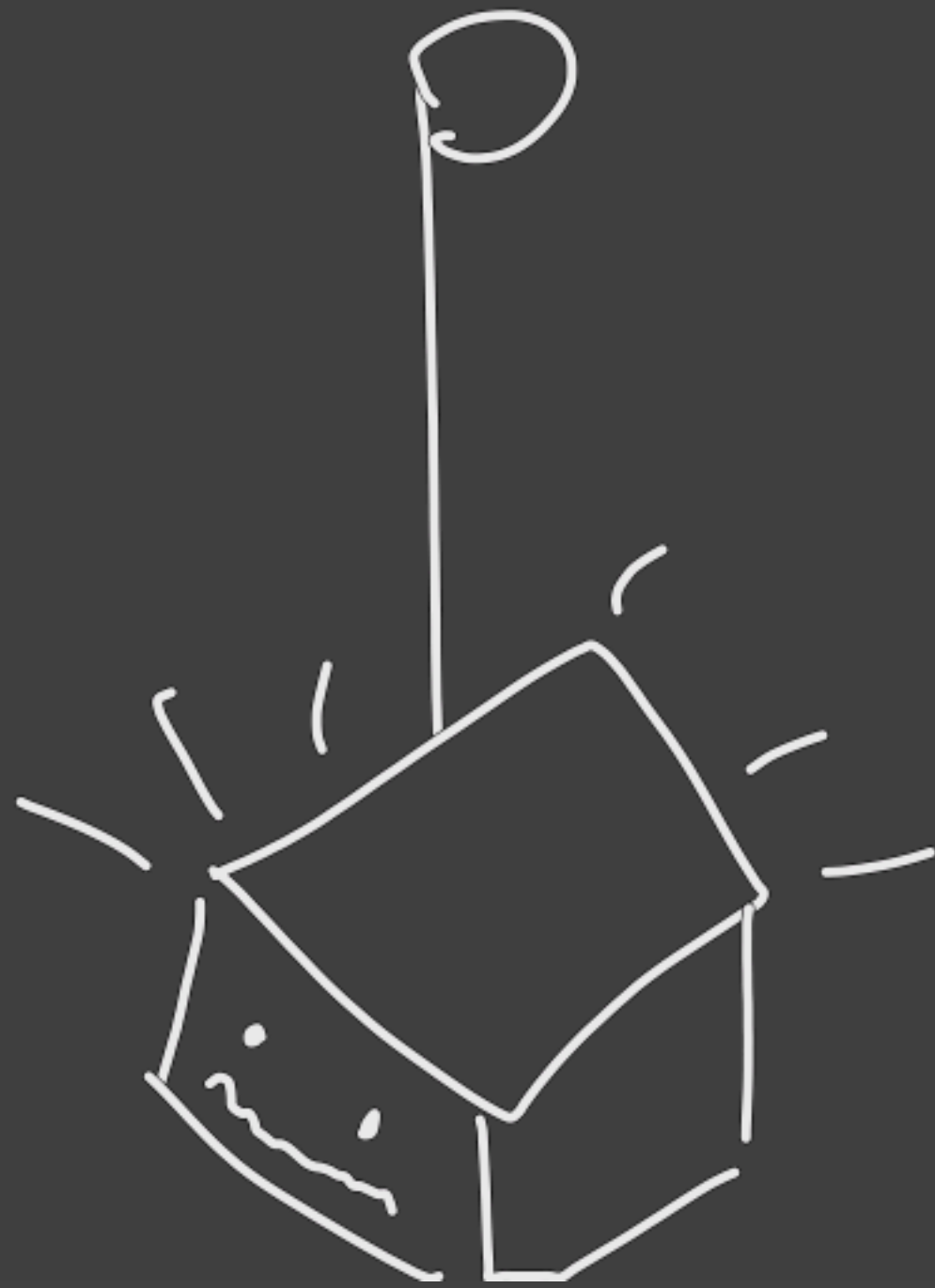
The other half use the other



Between subjects design

Half the participants use one interface

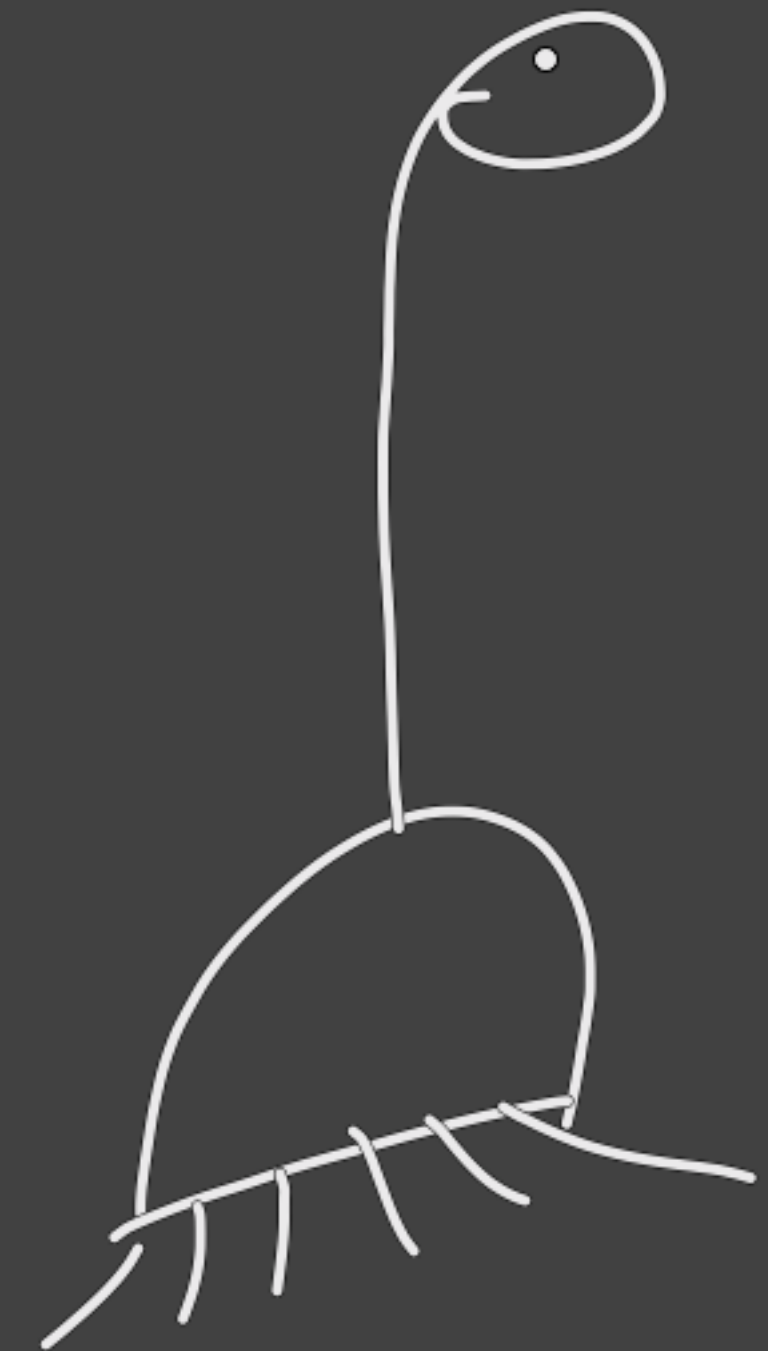
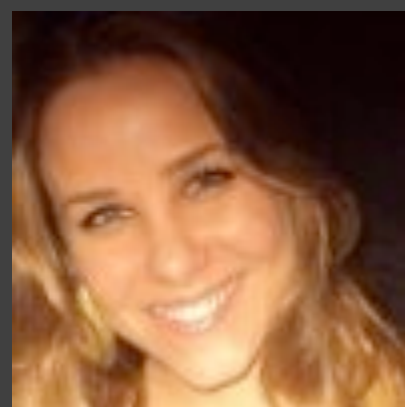
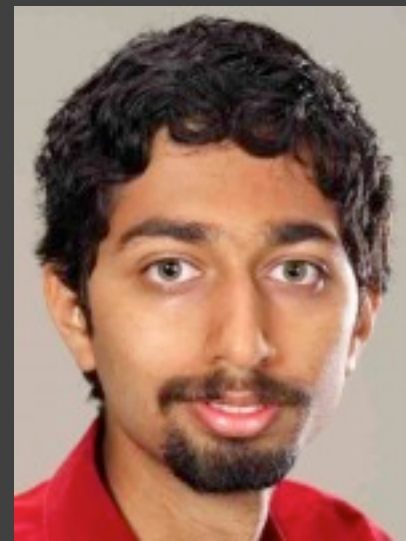
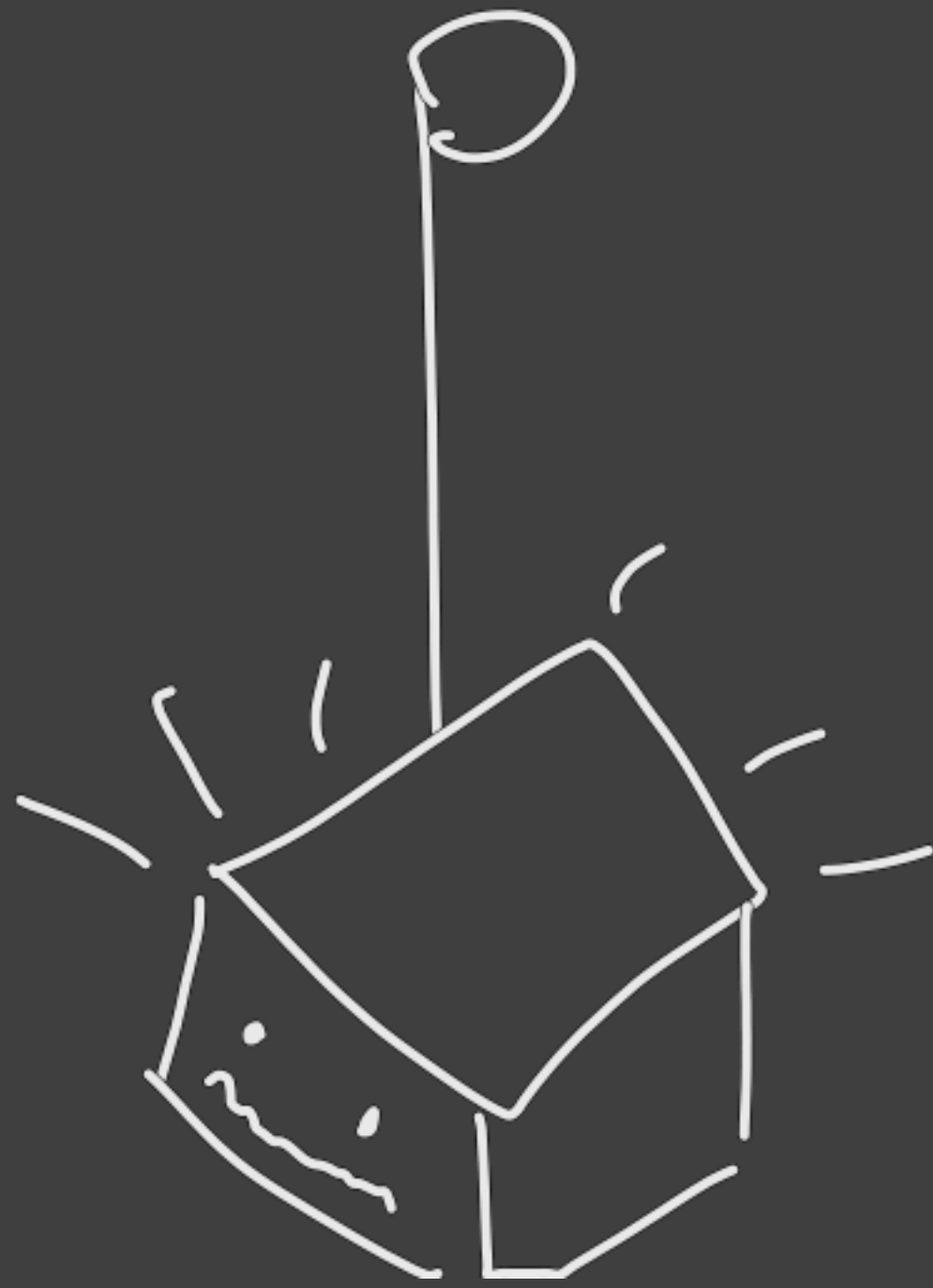
The other half use the other



Between subjects design

Half the participants use one interface

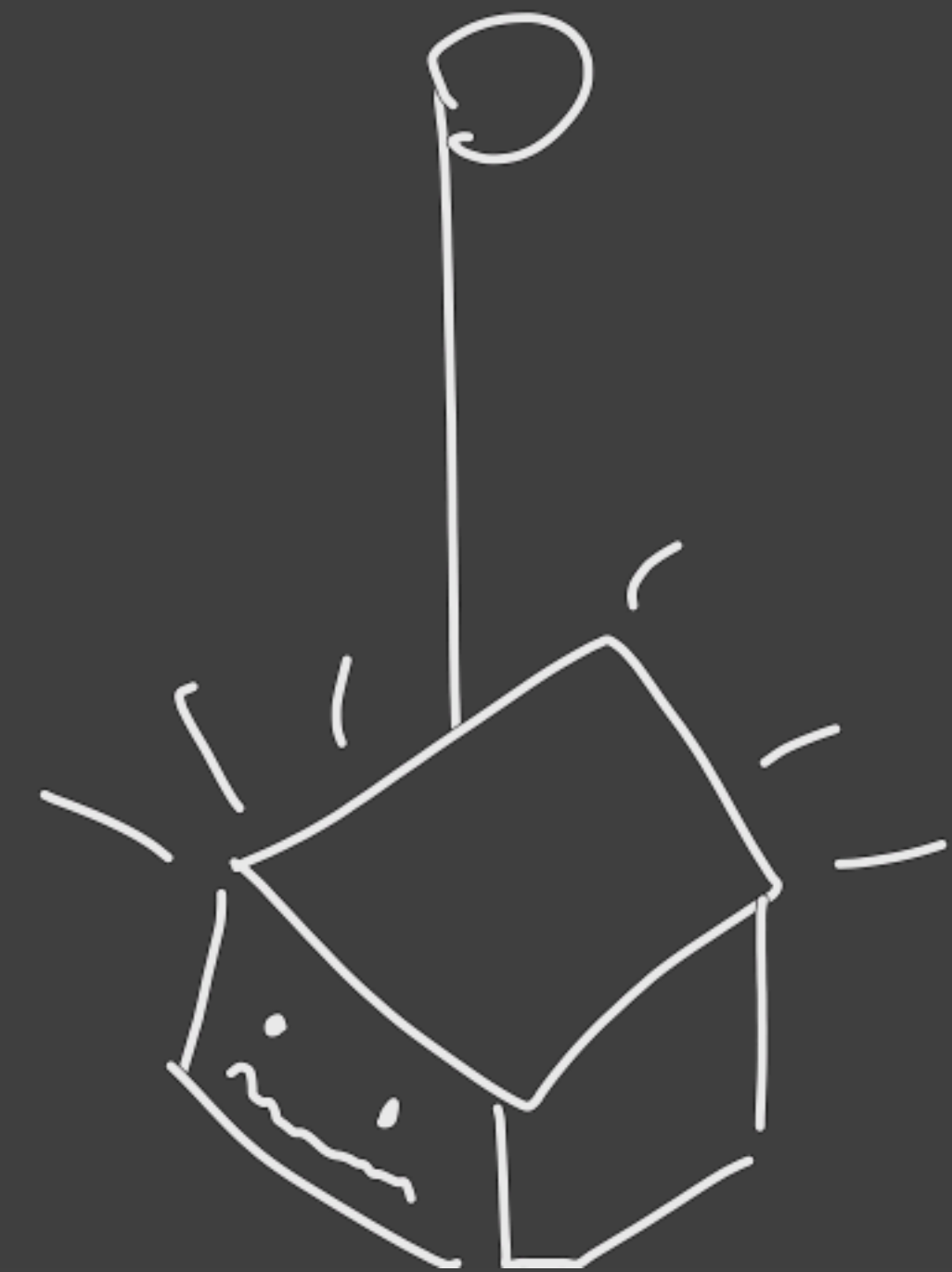
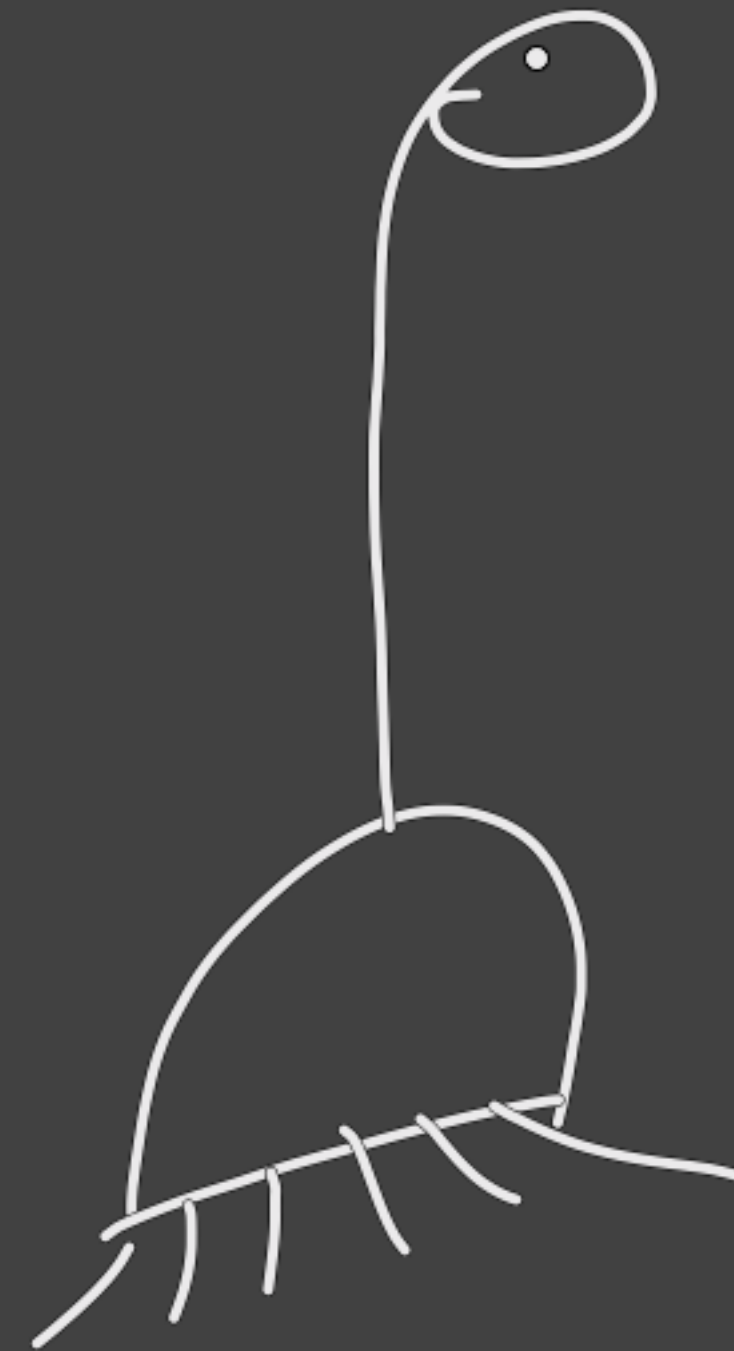
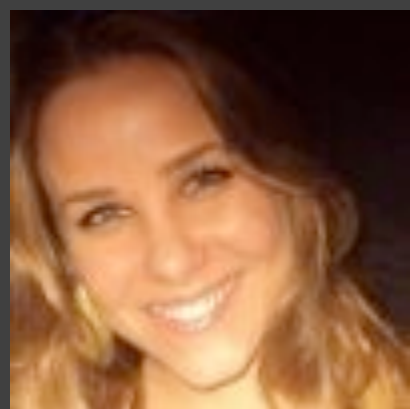
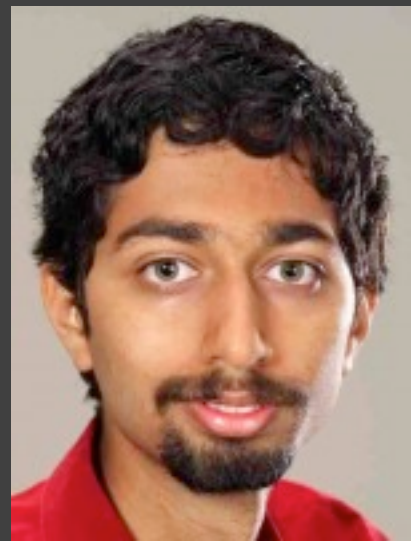
The other half use the other



Within subjects design

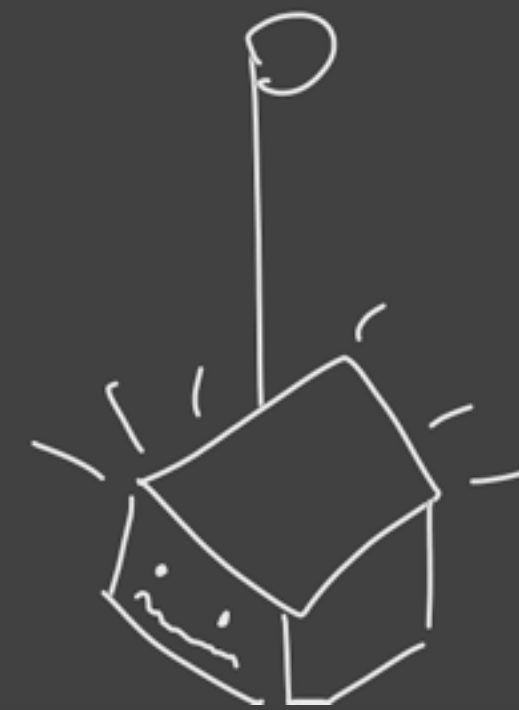
-

Everyone uses both interfaces

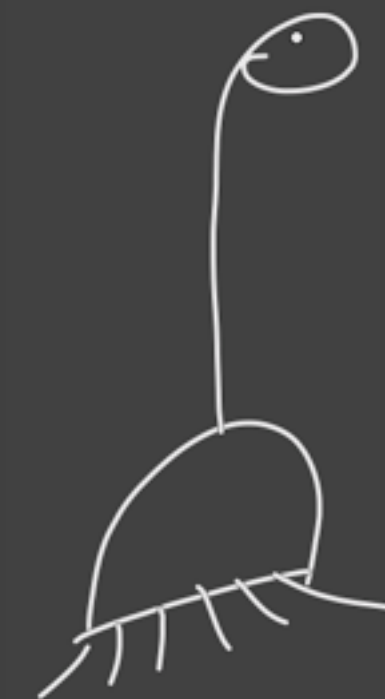
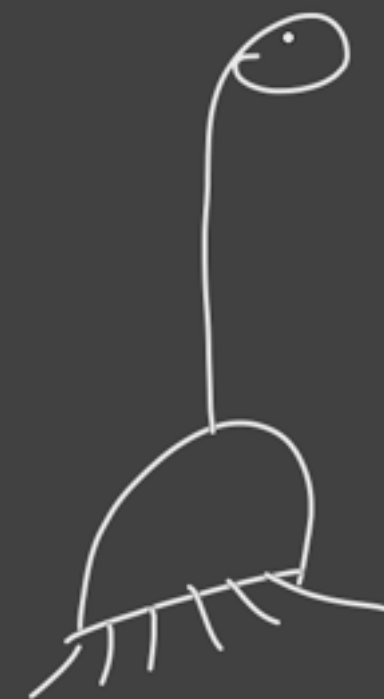
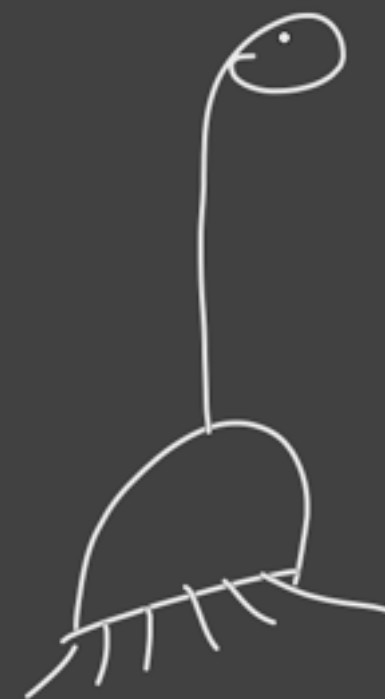
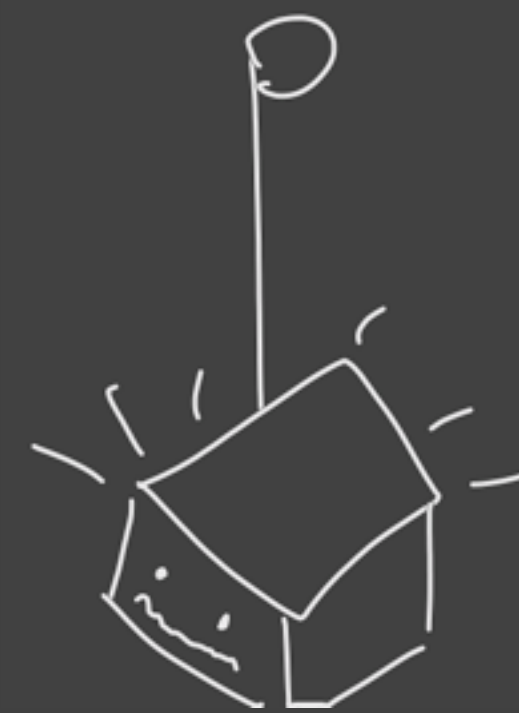
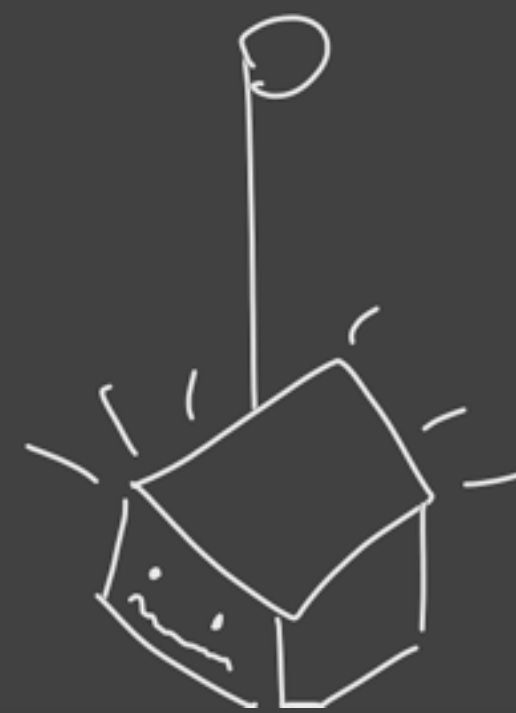
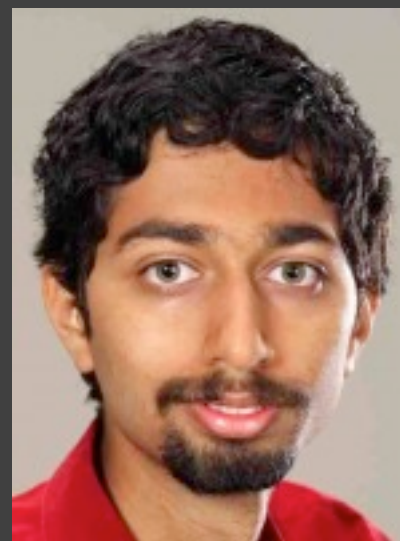


How Can We Address Ordering Effects?

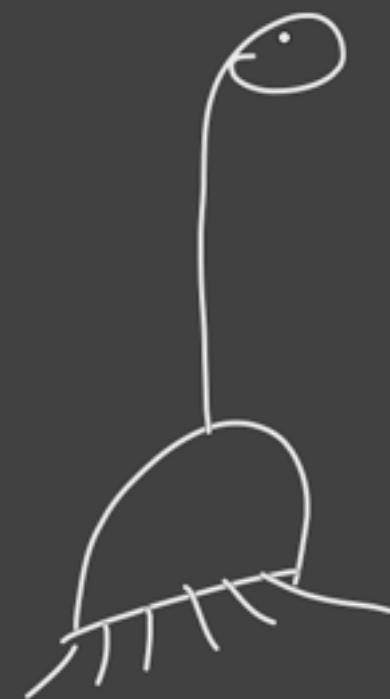
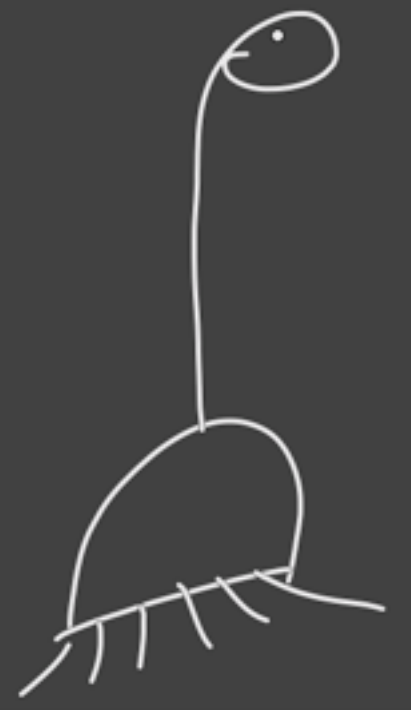
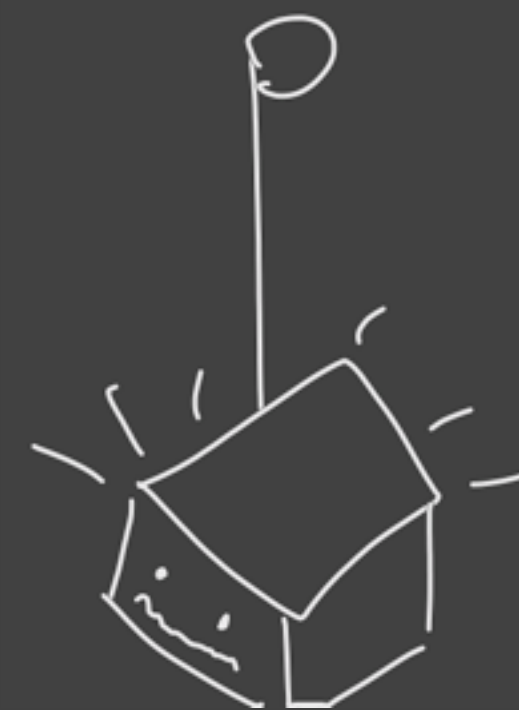
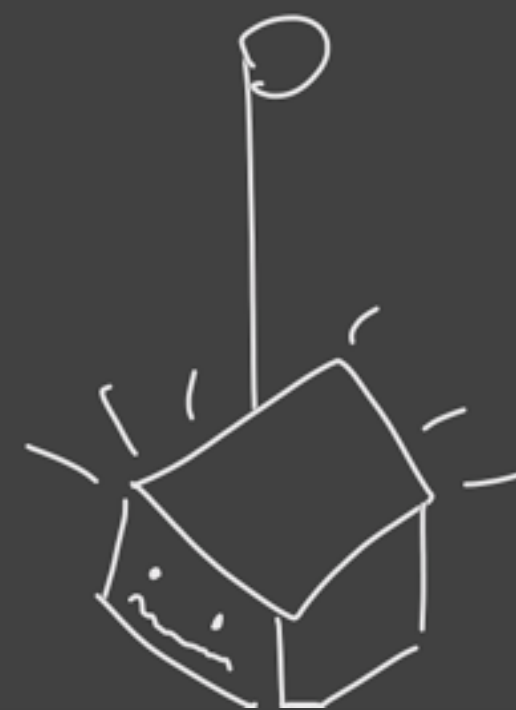
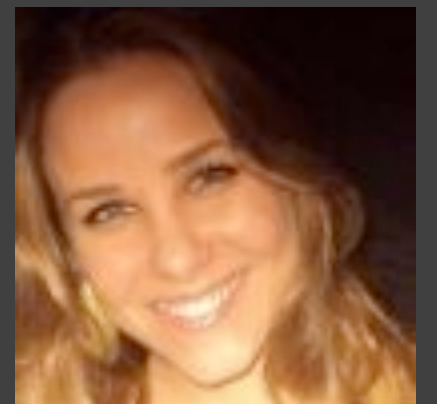
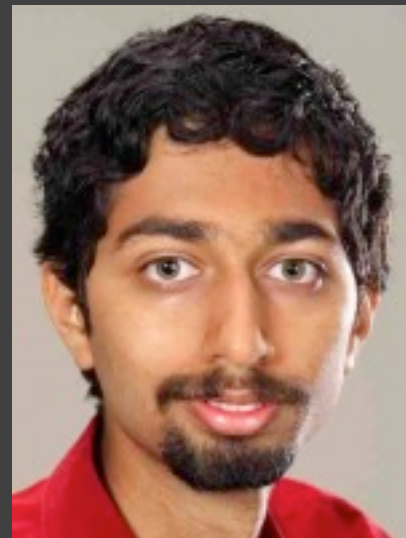
How Can We Address Ordering Effects?



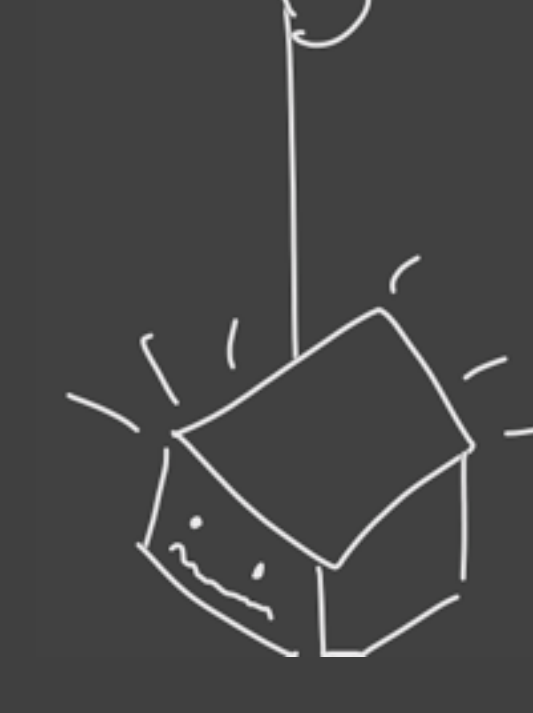
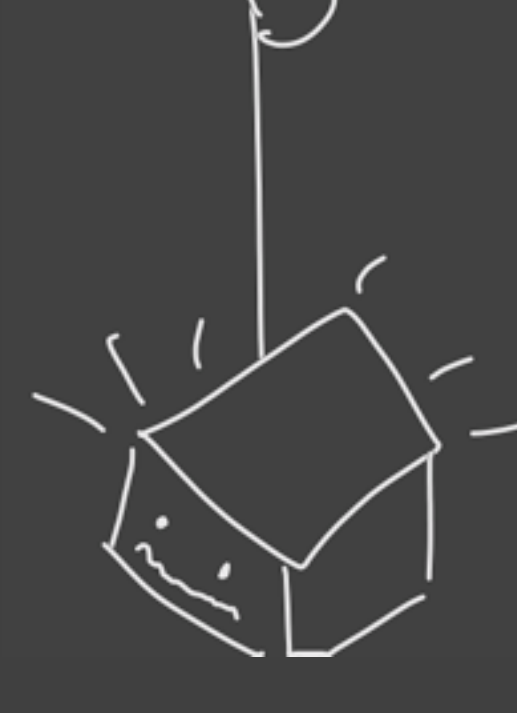
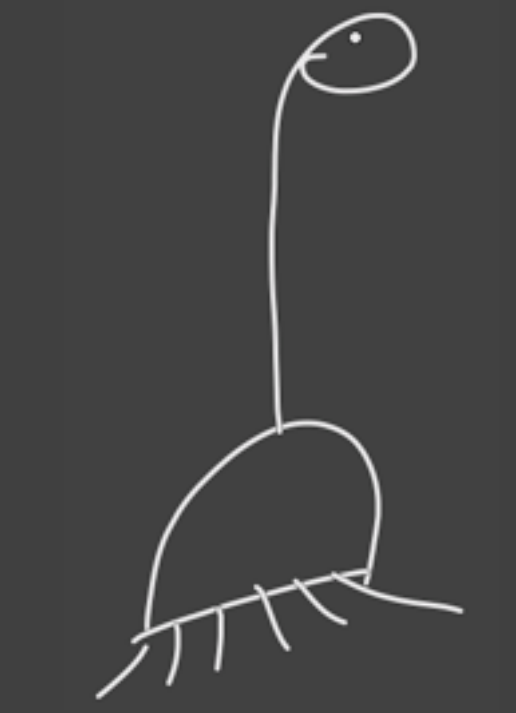
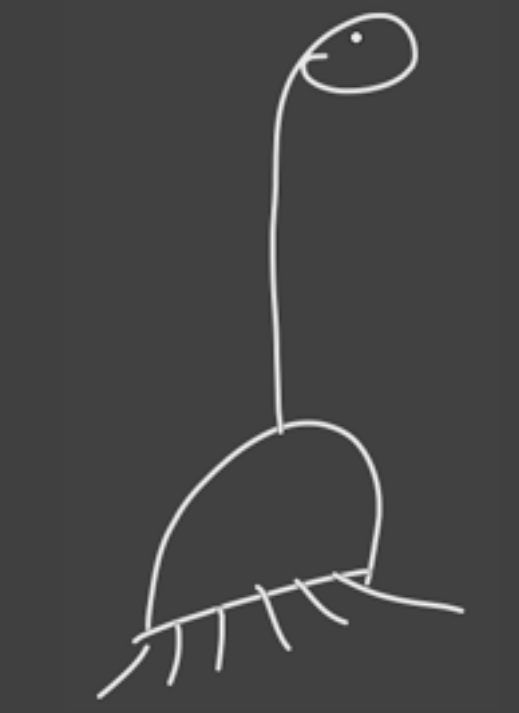
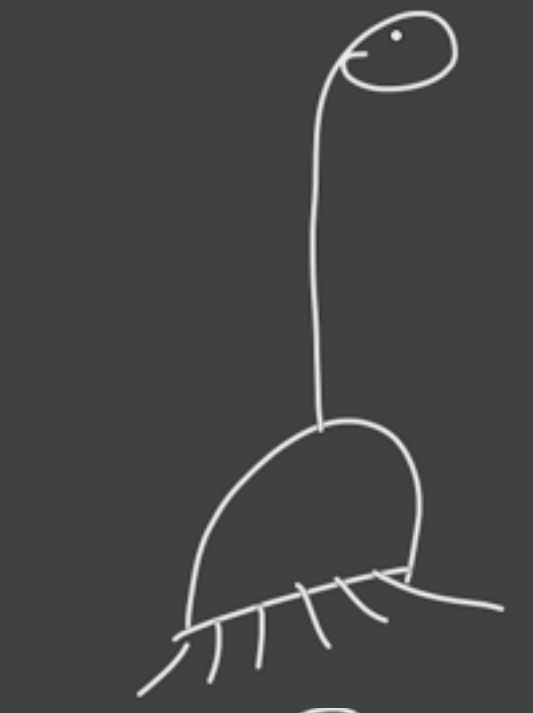
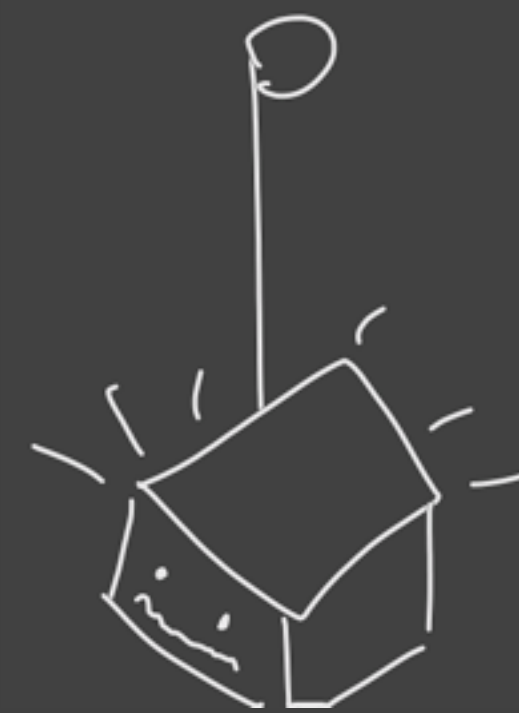
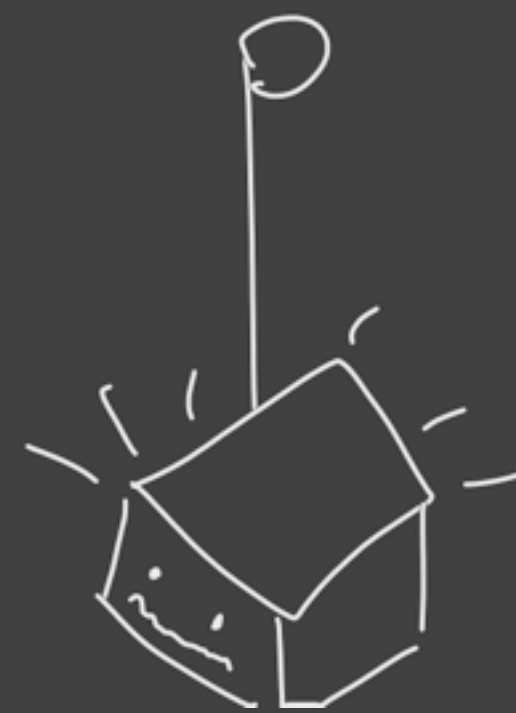
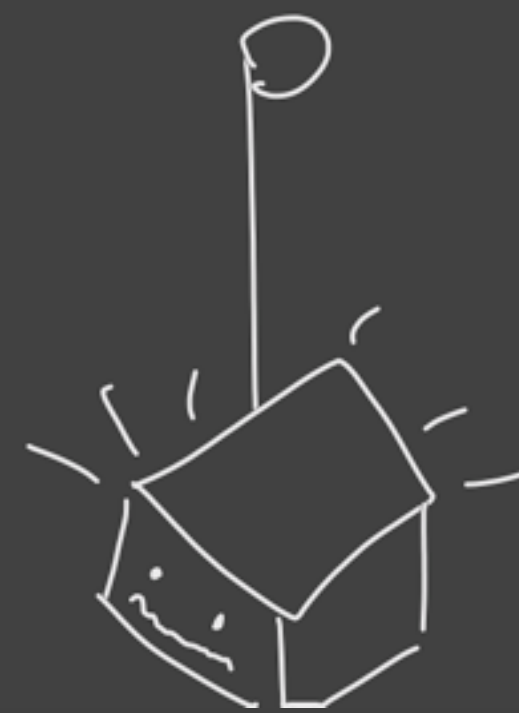
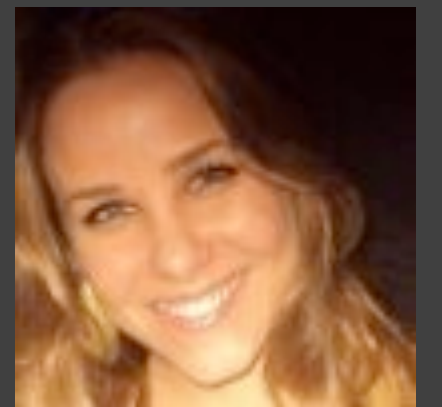
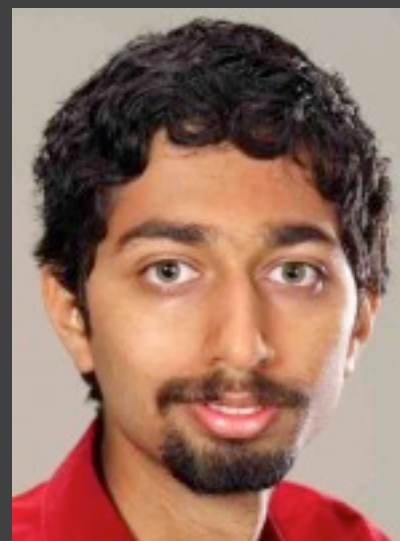
How Can We Address Ordering Effects?



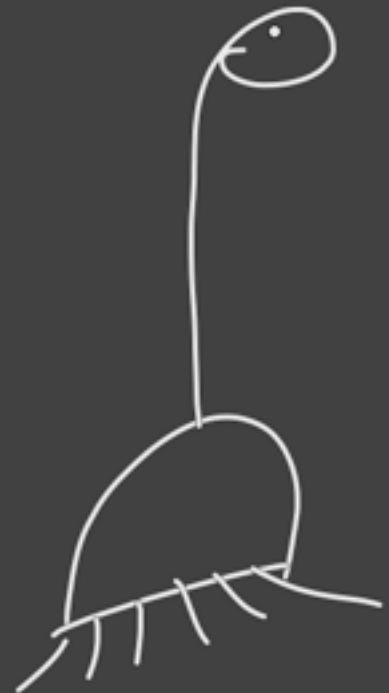
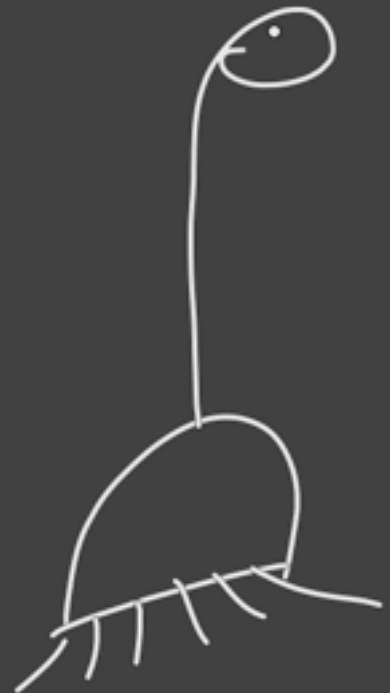
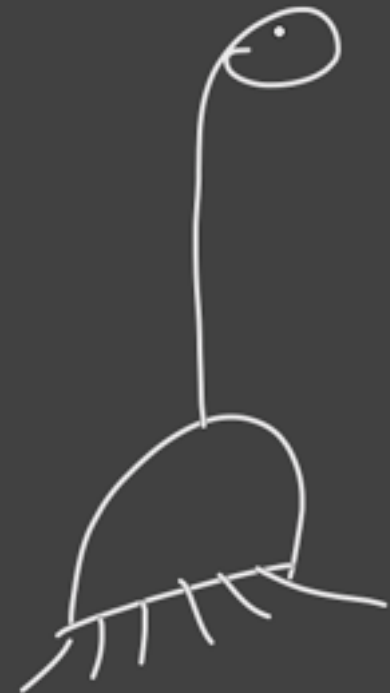
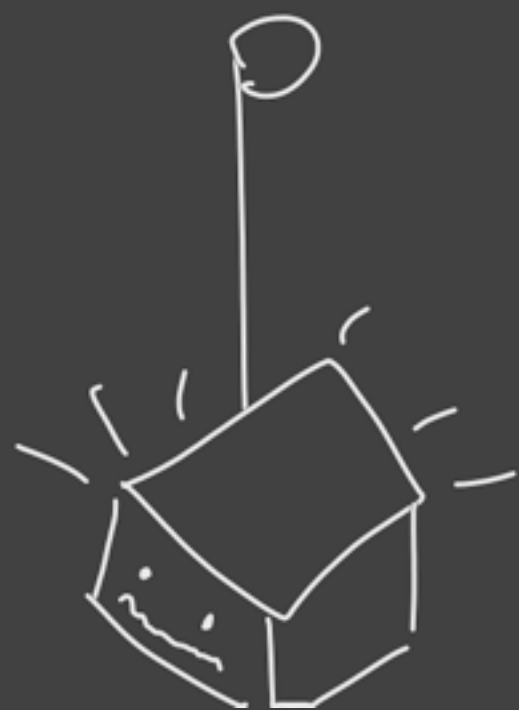
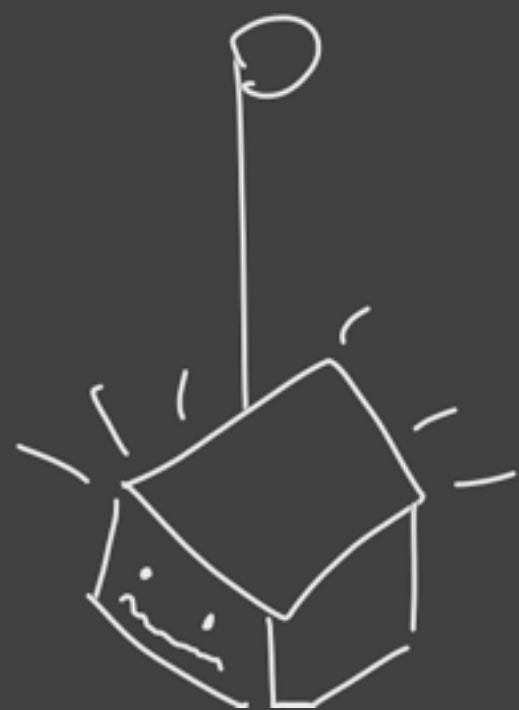
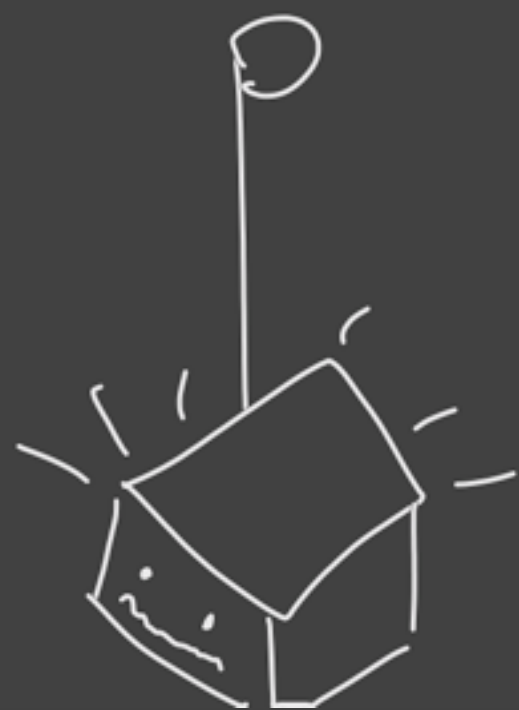
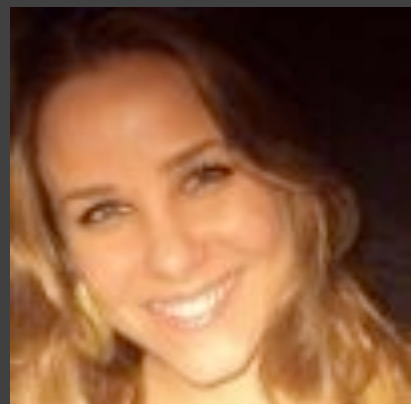
How Can We Address Ordering Effects?



How Can We Address Ordering Effects?









How Can We Address Ordering Effects?



How about individual differences?

What about for Three or More Alternatives?

Latin Square

 1	 2	3
 2	3	 1
3	 1	 2

The Importance of Random Assignment

The Importance of Random Assignment

- Typing in the morning versus the afternoon

The Importance of Random Assignment

The Importance of Random Assignment

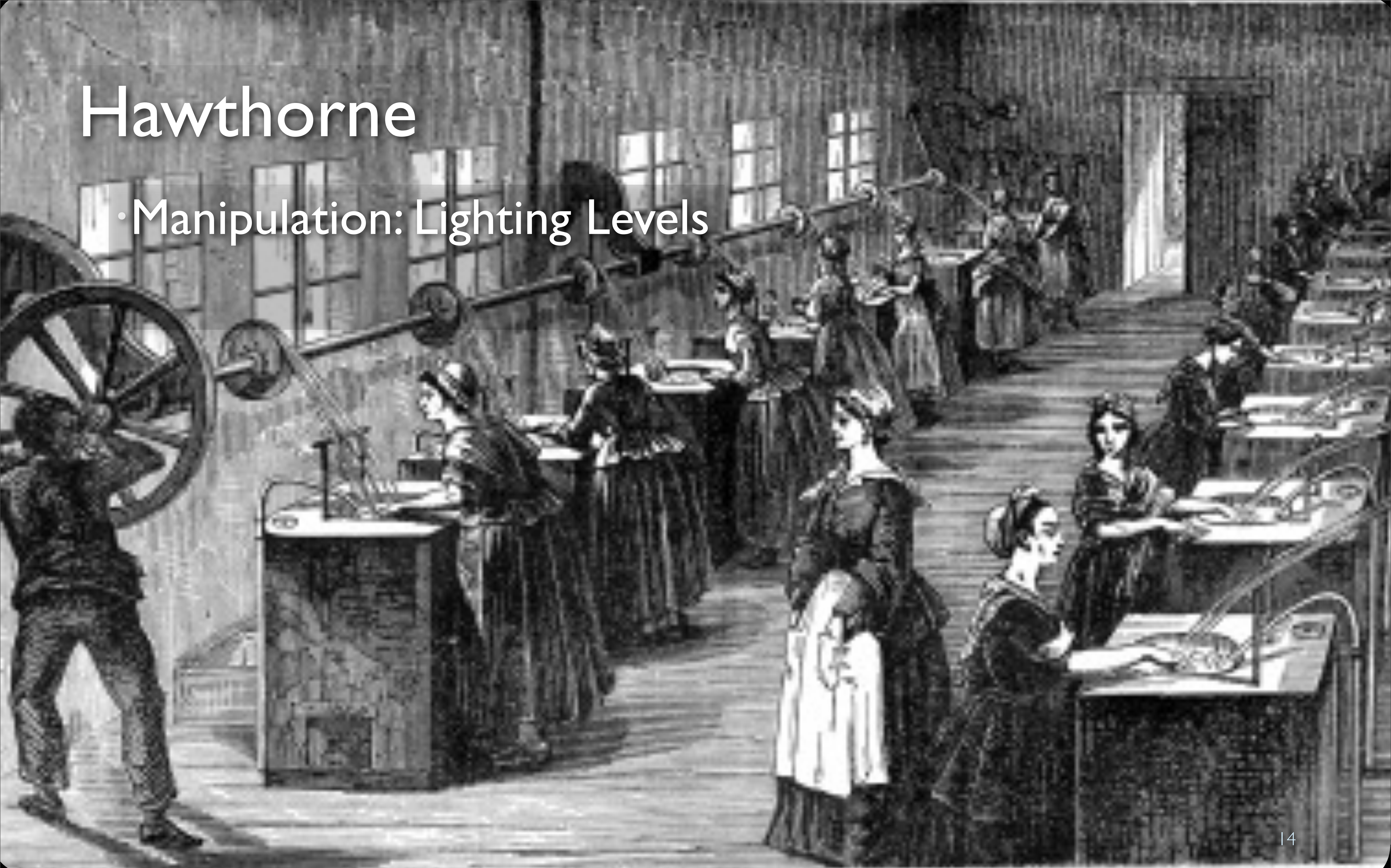
- Showing alternatives in sequence

Hawthorne



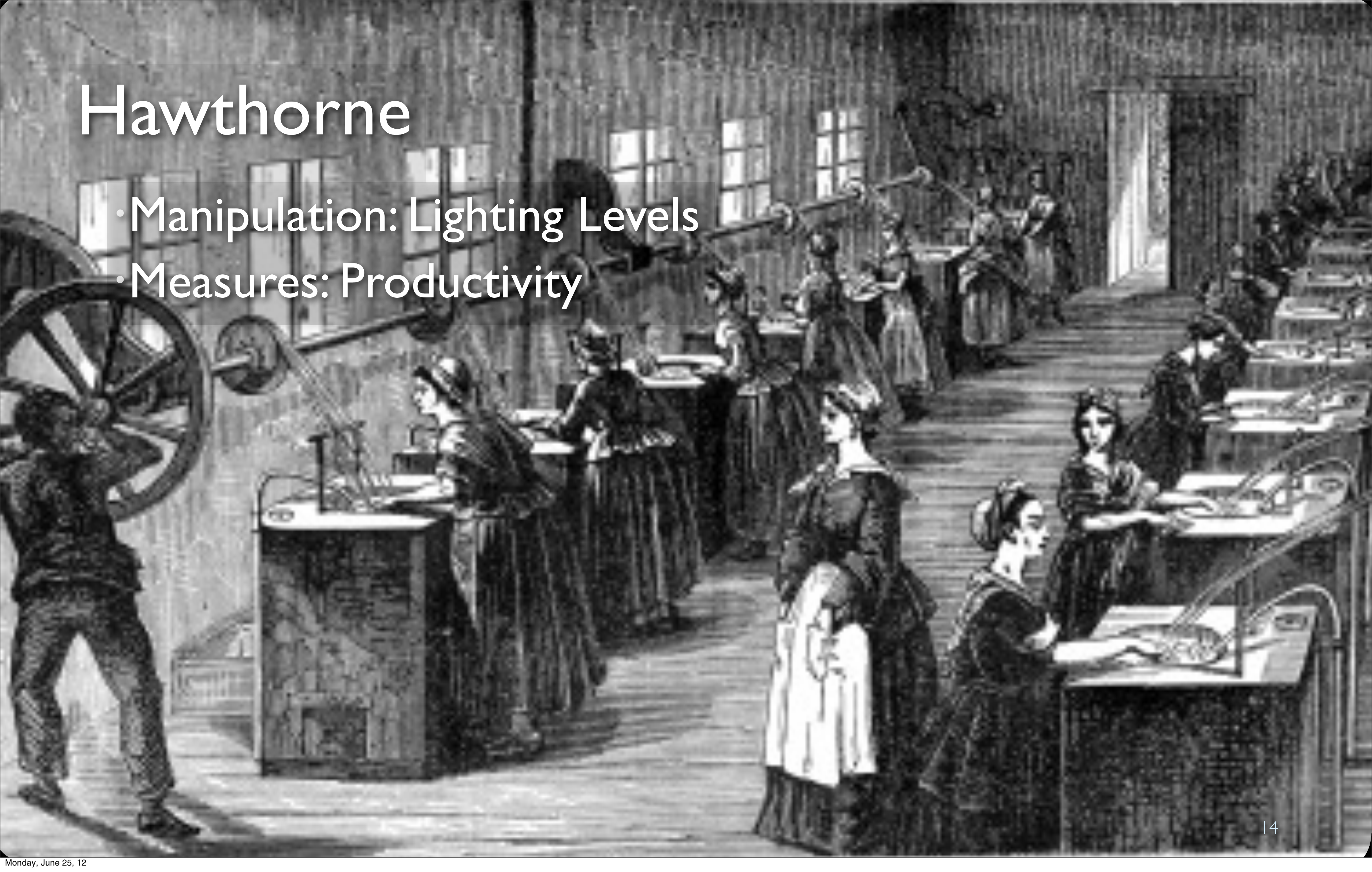
Hawthorne

- Manipulation: Lighting Levels



Hawthorne

- Manipulation: Lighting Levels
- Measures: Productivity



Counterbalanced Assignment

Counterbalanced Assignment

- Say you worry that typing speed will differentially affect interface usage (if you're e.g., designing a new spreadsheet)

Counterbalanced Assignment

- Say you worry that typing speed will differentially affect interface usage (if you're e.g., designing a new spreadsheet)
- You can use a pre-test to assign participants so that typing speed is roughly balanced across conditions

Counterbalanced Assignment

- Say you worry that typing speed will differentially affect interface usage (if you're e.g., designing a new spreadsheet)
- You can use a pre-test to assign participants so that typing speed is roughly balanced across conditions
- There are many techniques; the key is that each participant has an equal chance of landing in either condition

“Offline” Counterbalancing

TYPING SPEED

37

68

99

59

70

35

57

61

“Offline” Counterbalancing

- If you can pre-test everyone ahead of time, you can form matched pairs

TYPING SPEED

37

68

99

59

70

35

57

61

“Offline” Counterbalancing

- If you can pre-test everyone ahead of time, you can form matched pairs
- The assignment of each pair is the result of a coin toss

TYPING SPEED

37

68

99

59

70

35

57

61

“Offline” Counterbalancing

- If you can pre-test everyone ahead of time, you can form matched pairs
- The assignment of each pair is the result of a coin toss

TYPING SPEED	ORDERED
37	35
68	37
99	57
59	59
70	61
35	68
57	70
61	99

“Offline” Counterbalancing

- If you can pre-test everyone ahead of time, you can form matched pairs
- The assignment of each pair is the result of a coin toss

TYPING SPEED	ORDERED
37	{ 35
68	{ 37
99	{ 57
59	{ 59
70	{ 61
35	{ 68
57	{ 70
61	{ 99

“Online” Counterbalancing

TYPING SPEED

“Online” Counterbalancing

- If the pre-test happens online (*i.e.*, you can't pre-test everyone before the experiment begins)

TYPING SPEED

“Online” Counterbalancing

- If the pre-test happens online (*i.e.*, you can't pre-test everyone before the experiment begins)
- Pick some threshold that's likely to be about the middle

TYPING SPEED

“Online” Counterbalancing

- If the pre-test happens online (*i.e.*, you can’t pre-test everyone before the experiment begins)
- Pick some threshold that’s likely to be about the middle

TYPING SPEED

35

40

90

68

...

- I think of pre-test-like counterbalancing as helping the law of large numbers work a little bit faster

A danger: regression

A danger: regression

- Let's find heady coins

A danger: regression

- Let's find heady coins
- First, let's flip all the coins (our pre-test)

A danger: regression

- Let's find heady coins
- First, let's flip all the coins (our pre-test)
- If they land heads more than half, we'll call them heady

A danger: regression

- Let's find heady coins
- First, let's flip all the coins (our pre-test)
- If they land heads more than half, we'll call them heady
- Now let's feed them a snack

A danger: regression

- Let's find heady coins
- First, let's flip all the coins (our pre-test)
- If they land heads more than half, we'll call them heady
- Now let's feed them a snack
- Does snacking increase the natural tendency of coins?

15
15
15
14
14
13
13
12
12
12
12
12
12
11
11
10
10
10
10
10
9
9
9
8
8
8
7
6
6
5

15
15
15
14
14
13
13
12
12
12
12
12
12
11
11
10
10
10
10
10
9
9
9
8
8
8
7
6
6
5

10
14
8
8
15
12
12
8
10
12
10
13
8
9
11
8
5
11
15
10
7
9
9
10
7
8
10
12
14
9

Avoiding the Danger

Avoiding the Danger

- If the pre-test is used to counterbalance, and assignment is random, then the error goes away

Should every participant use
every alternative?

Three Major Strategies

Three Major Strategies

- **Within-subjects:** everyone tries all the options. Good when you're not worried about learning/practice/exposure issues (that trying one version will 'pollute' the data from another version)

Three Major Strategies

- **Within-subjects:** everyone tries all the options. Good when you're not worried about learning/practice/exposure issues (that trying one version will 'pollute' the data from another version)
- **Between-participants:** each person tries one. Requires more people, and more attention to fair assignment. Has the benefit that each participant is uncorrupted (at least by the study...)

Three Major Strategies

- **Within-subjects:** everyone tries all the options. Good when you're not worried about learning/practice/exposure issues (that trying one version will 'pollute' the data from another version)
- **Between-participants:** each person tries one. Requires more people, and more attention to fair assignment. Has the benefit that each participant is uncorrupted (at least by the study...)
- **Counterbalancing** can help minimize variation in a between-subjects design

Further Reading

Further Reading

- Martin, Doing Psychology Experiments