CSE440: Introduction to HCI

Methods for Design, Prototyping and Evaluating User Interaction

Lecture 11:
Testing

Nigini Oliveira
Abhinav Yadav
Liang He
Angel Vuong
Jeremy Viny

DESIGN
USE
BUILD

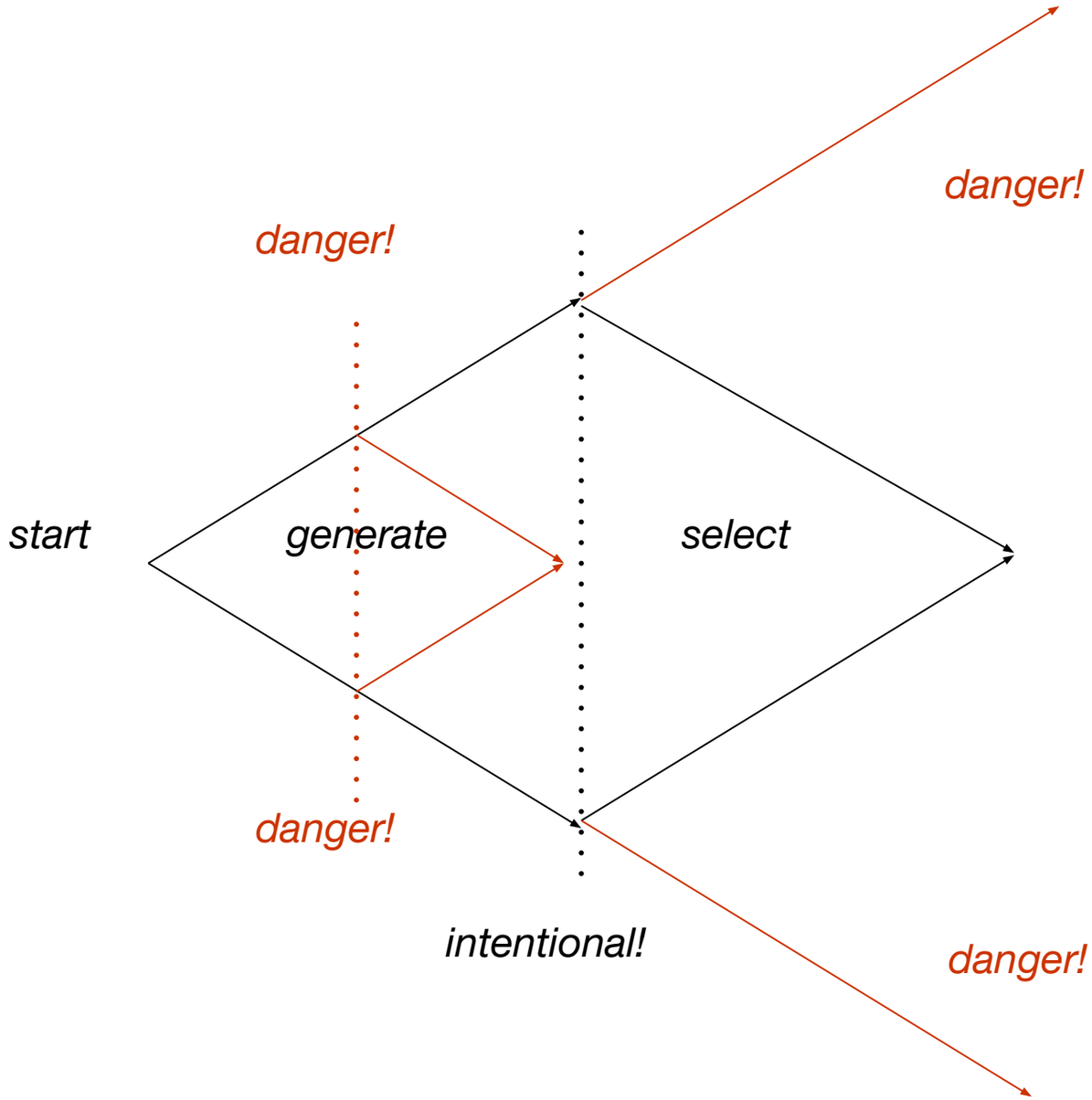
PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

What we will do today

Design sprint

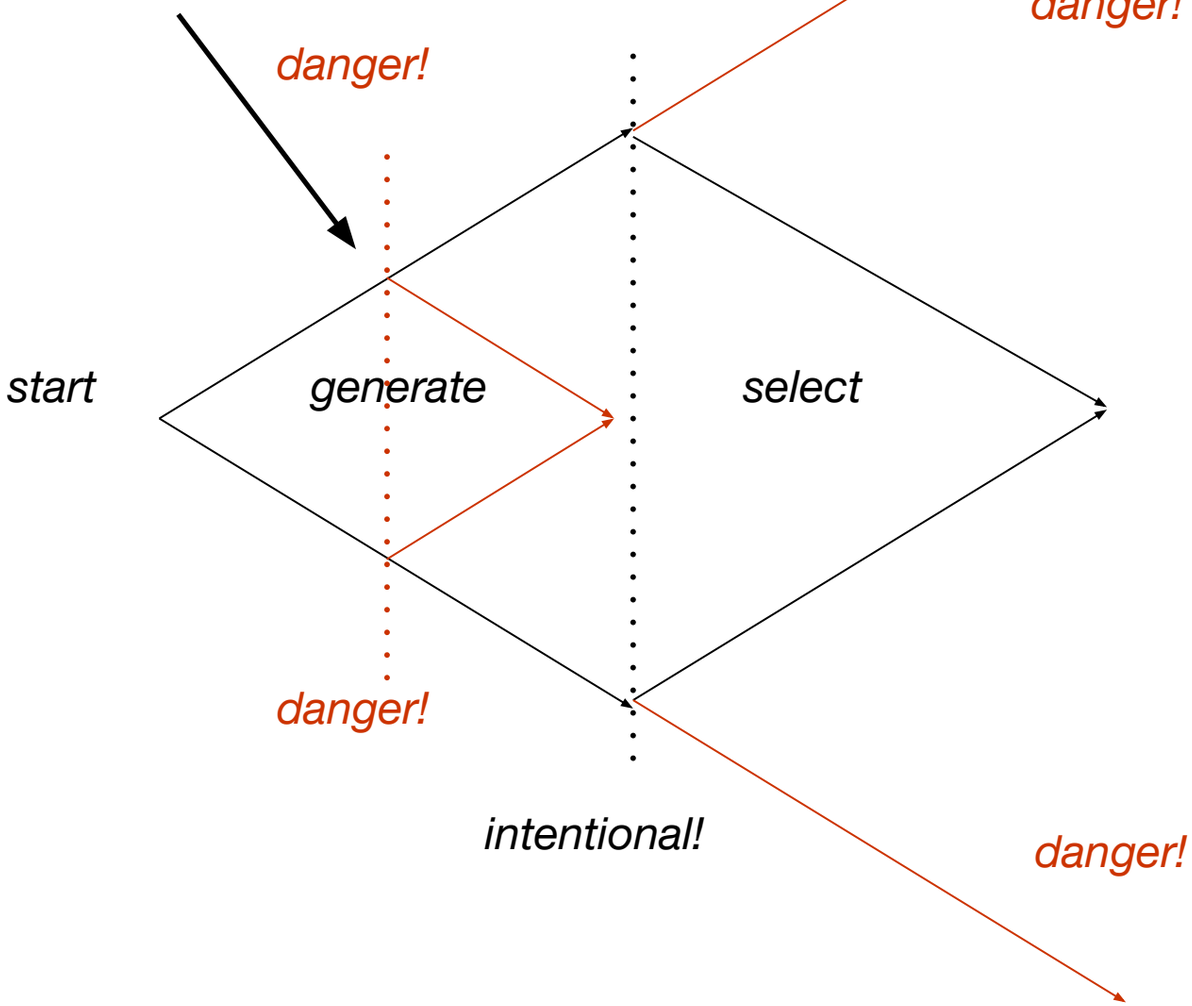
Testing (beyond paper prototypes)

But first a reminder

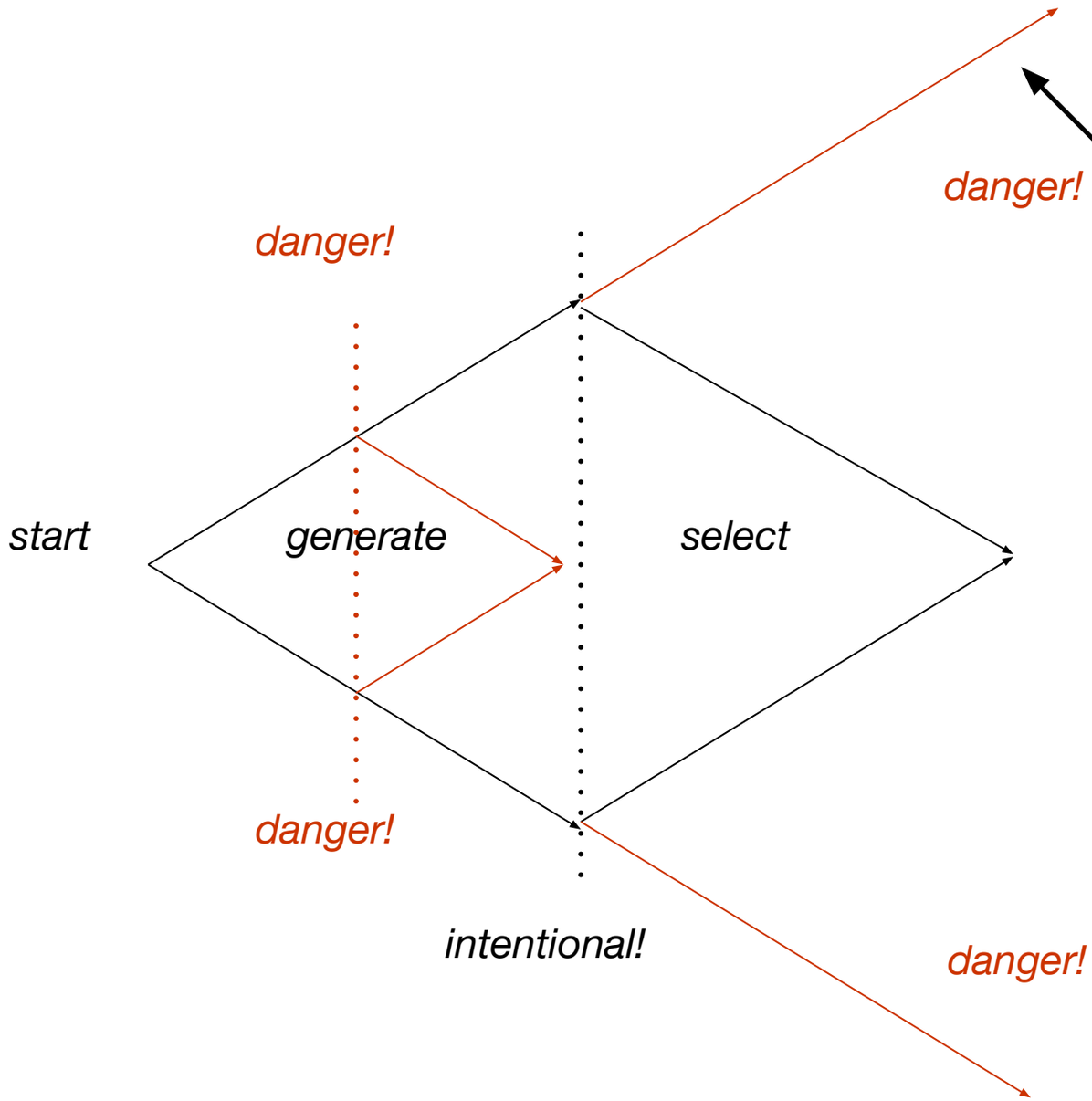


But first a reminder

Some of you decided on one idea too fast!

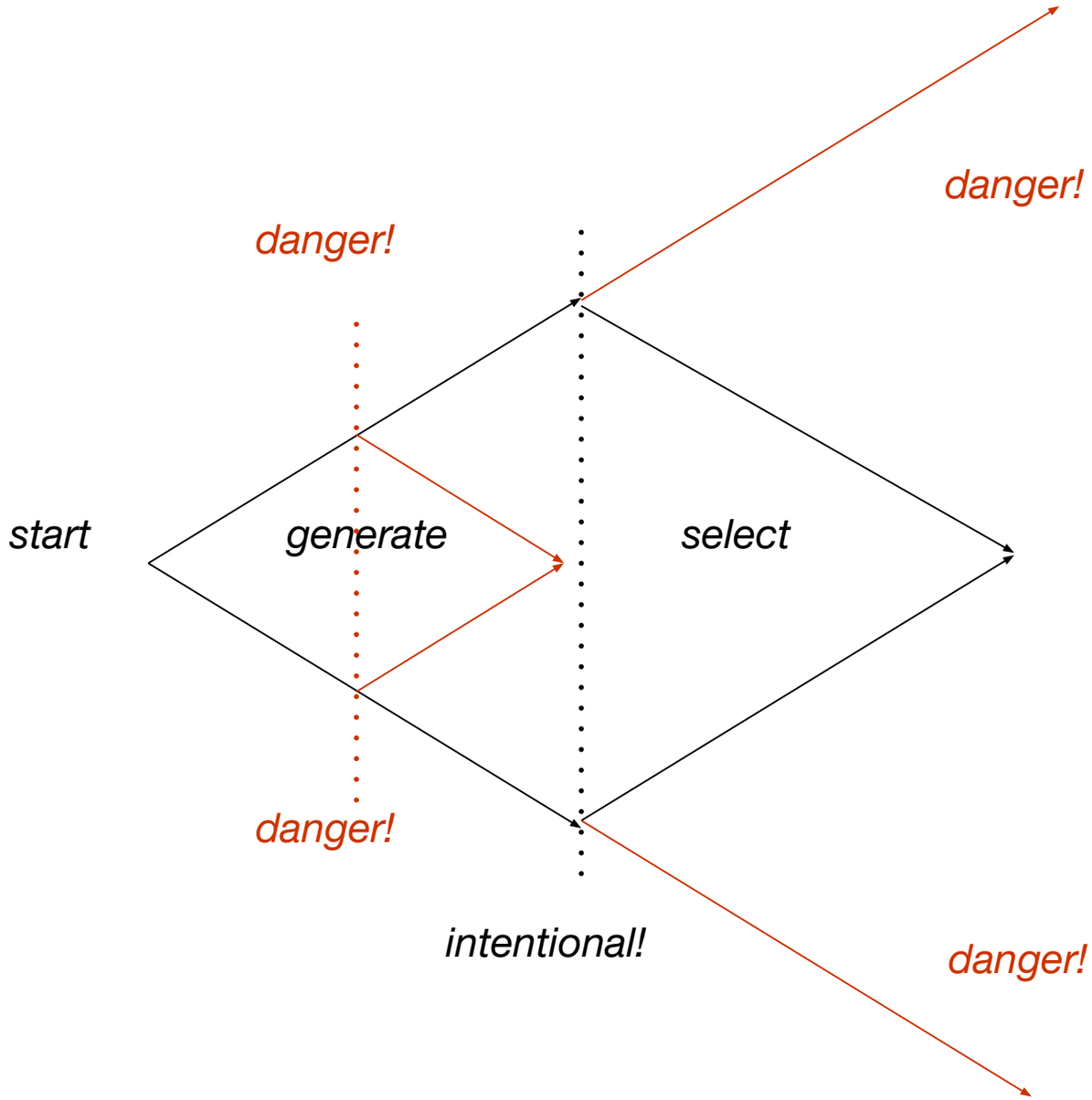


But first a reminder



Some of you are not deciding on one idea fast enough

But first a reminder



Make sure to meet in your team and decide!

You still have time to adjust/explain things in your **final report**.

Upcoming

May 6	May 7	May 8	May 9	May 10
Nigini at CHI	Nigini at CHI	Nigini at CHI	Nigini at CHI	Nigini at CHI
	Testing 10:00 - 11:20 | OUG 136	2g - Design Review	Heuristic Evaluation 10:00 - 11:20 | OUG 136	Section 10:30 - 11:20 | MGH 058 11:30 - 12:20 | MGH 058 1:30 - 2:20 | MGH 058 2:30 - 3:20 | MGH 058
	Reading 2: Storyboarding and Video Prototyping			2h - Getting the Right Design

May 13	May 14	May 15	May 16	May 17
3a - Paper Protoype	Accessibility 10:00 - 11:20 | OUG 136	3b - Heuristic Evaluation	Patterns 10:00 - 11:20 | OUG 136	Section 10:30 - 11:20 | MGH 058 11:30 - 12:20 | MGH 058 1:30 - 2:20 | MGH 058 2:30 - 3:20 | MGH 058
		Nigini's office hours 10:00 - 12:00 Allen Center 338		Reading 3: Usability Testing
				3c - Usability Testing Check-In

Design Sprint

Design Sprint (30 minutes)

Form groups of 3
(find at least one member who is not usually in your usual groups)

Read through the handout

5 minutes: Sketch out a few designs and decide on one

15 minutes: Prototype your design

5 minutes: Test your design with one of the TAs, myself, or with a member of another group

Design Sprint (30 minutes)

Form groups of 3
(find at least one member who is not usually in your usual groups)

Read through the handout

5 minutes: Sketch out a few designs and decide on one

15 minutes: Prototype your design

5 minutes: Test your design with one of the TAs, myself, or with a member of another group

Design Sprint (30 minutes)

Form groups of 3
(find at least one member who is not usually in your usual groups)

Read through the handout

5 minutes: Sketch out a few designs and decide on one

15 minutes: Prototype your design

5 minutes: Test your design with one of the TAs, myself, or with a member of another group

Design Sprint (30 minutes)

Form groups of 3
(find at least one member who is not usually in your usual groups)

Read through the handout

5 minutes: Sketch out a few designs and decide on one

15 minutes: Prototype your design

5 minutes: Test your design with one of the TAs, myself, or with a member of another group

Design Sprint (30 minutes)

Form groups of 3
(find at least one member who is not usually in your usual groups)

Read through the handout

5 minutes: Sketch out a few designs and decide on one

15 minutes: Prototype your design

5 minutes: Test your design with one of the TAs, myself, or with a member of another group

Discussion

What worked well during the paper prototyping?

What did you learn about coffee ordering?

How would you change your design?

What did you come up with?

Testing

Testing

User tests vs. Usability tests

Different ways for gaining insights

- Quick and dirty: informal discussions with users
  sometimes done using paper prototypes/Wizard-of-Oz testing
- Usability tests in lab
  e.g., Human Task Performance Measures
- Remote usability tests
  e.g., using recording software
- A/B testing

Remote Usability Testing

Conferencing-based testing

Use tools like video conferencing, instant messaging, and screencasting to test with a remote participant

Unmoderated remote testing

Automatic logging and some analysis of usage

Controlled online A/B experiments

Carefully measure results of showing different versions to different sets of live customers

Unmoderated Remote Usability Test

Move usability testing online

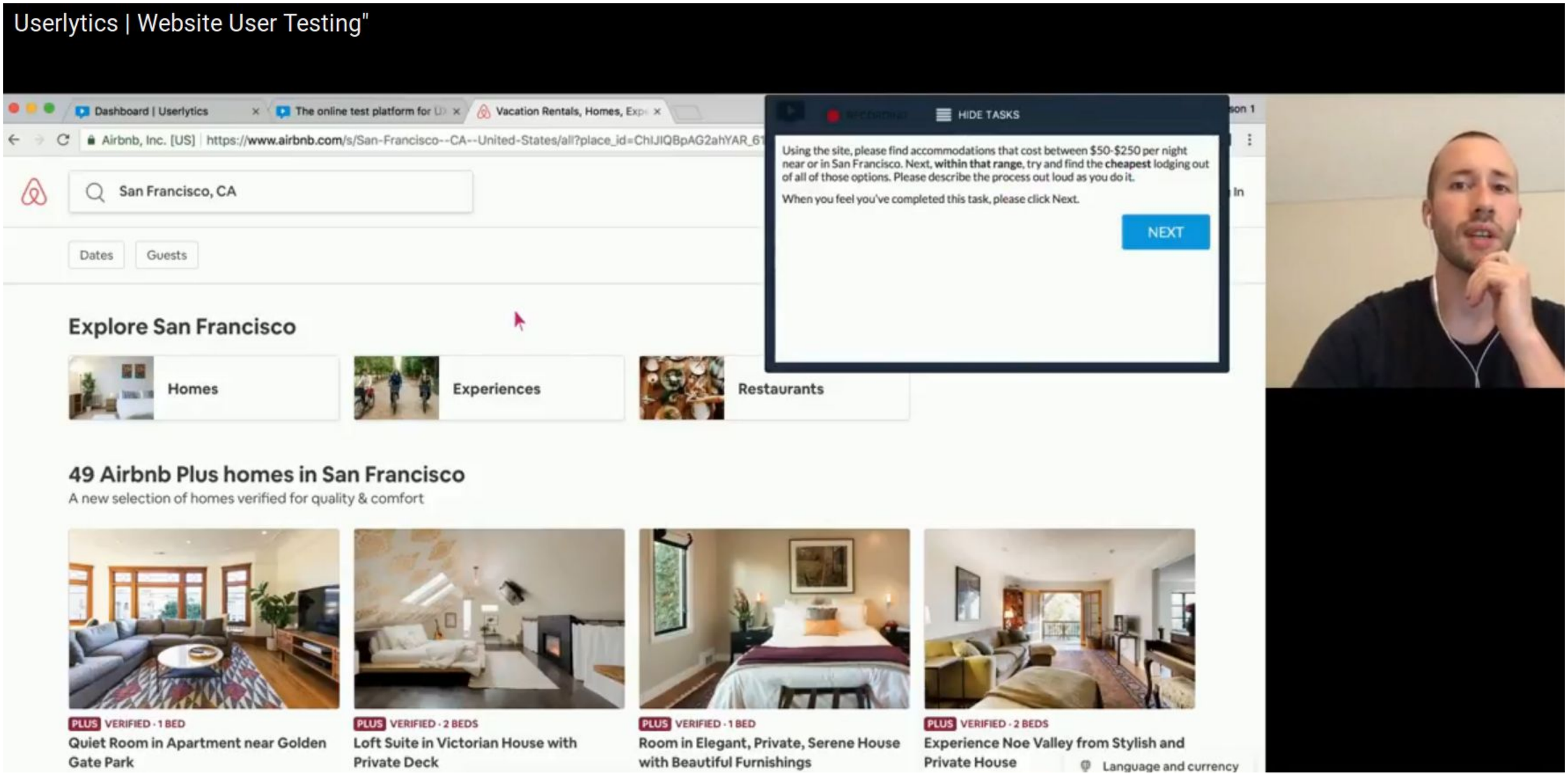
- participants access the “lab” via web
- answer questions & complete tasks in “survey”
- records actions or screens for playback
- can test many people completing many tasks

Analyze data individually or in aggregate

- playback individual sessions
- find general problem areas
- if needed, look more closely with traditional methods

More: https://www.nngroup.com/articles/unmoderated-user-testing-tools/

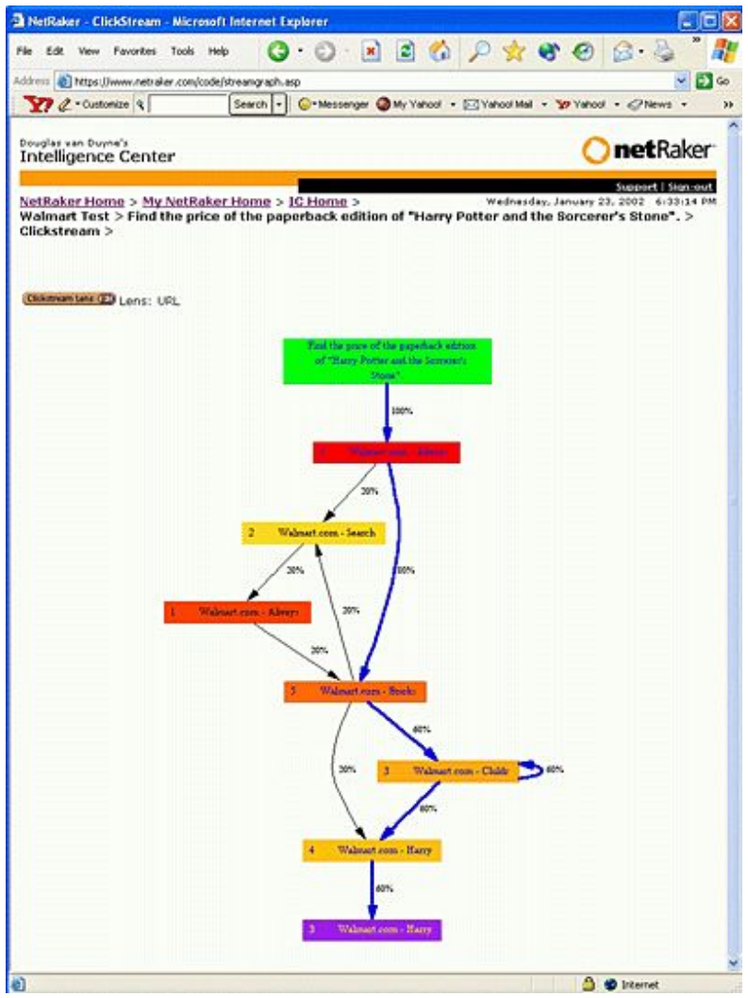
Unmoderated Remote Usability Test



Unmoderated Remote Usability Test

https://www.loop11.com/

Unmoderated Remote Usability Test



Controlled A/B Experiments

Many names for concept

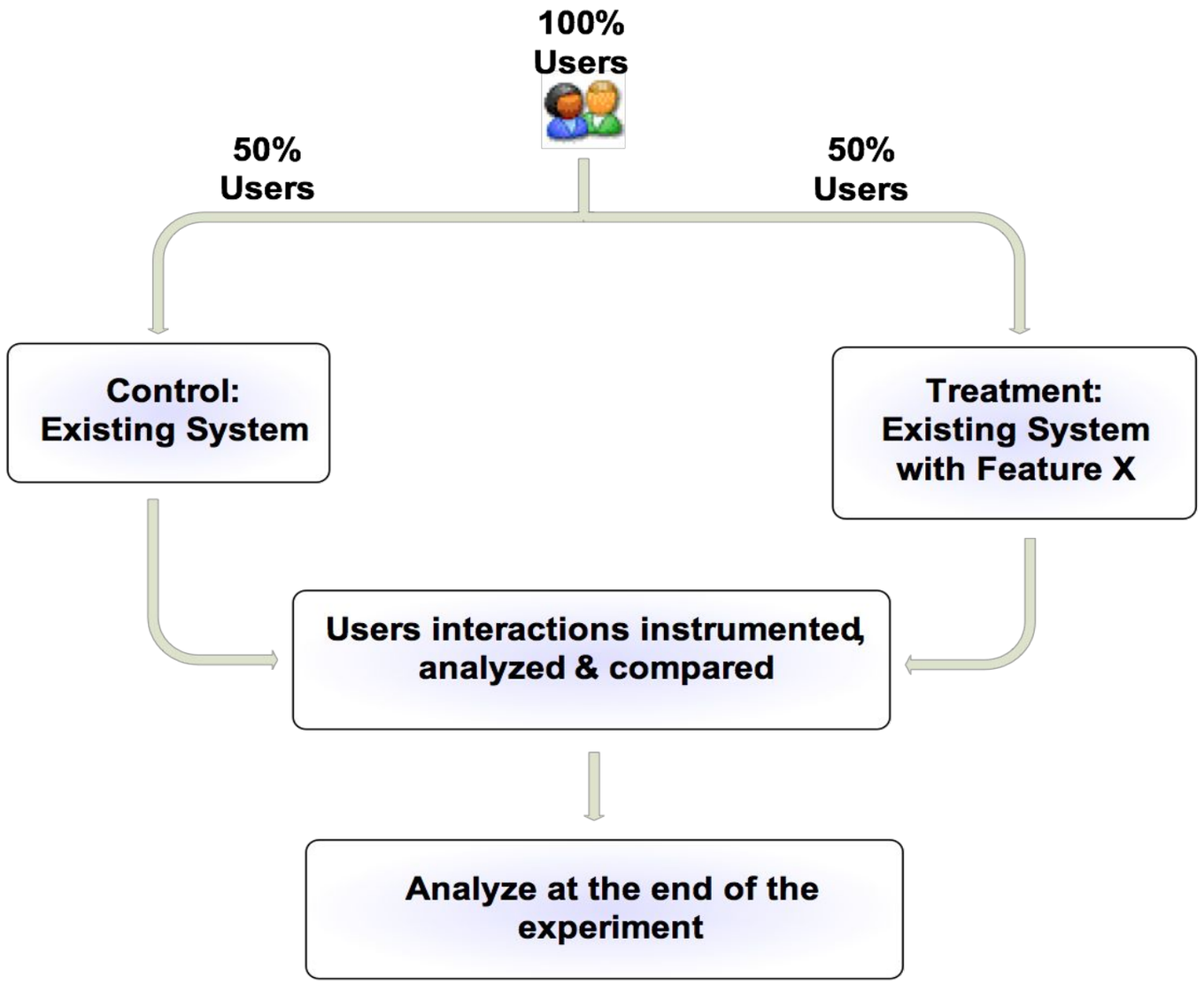
A/B tests or
Control/Treatment

Randomized
Experimental Design

Controlled experiments

Split testing

Parallel flights



Controlled A/B Experiments

Example: Amazon Shopping Cart Recommendations

- Add an item to your shopping cart
- Most sites show the cart

At Amazon, Greg Linden had idea to show recommendations based on cart items

Controlled A/B Experiments

Evaluation

- Pro: cross-sell more items
- Con: distract people from checking out

Highest Paid Person's Opinion:

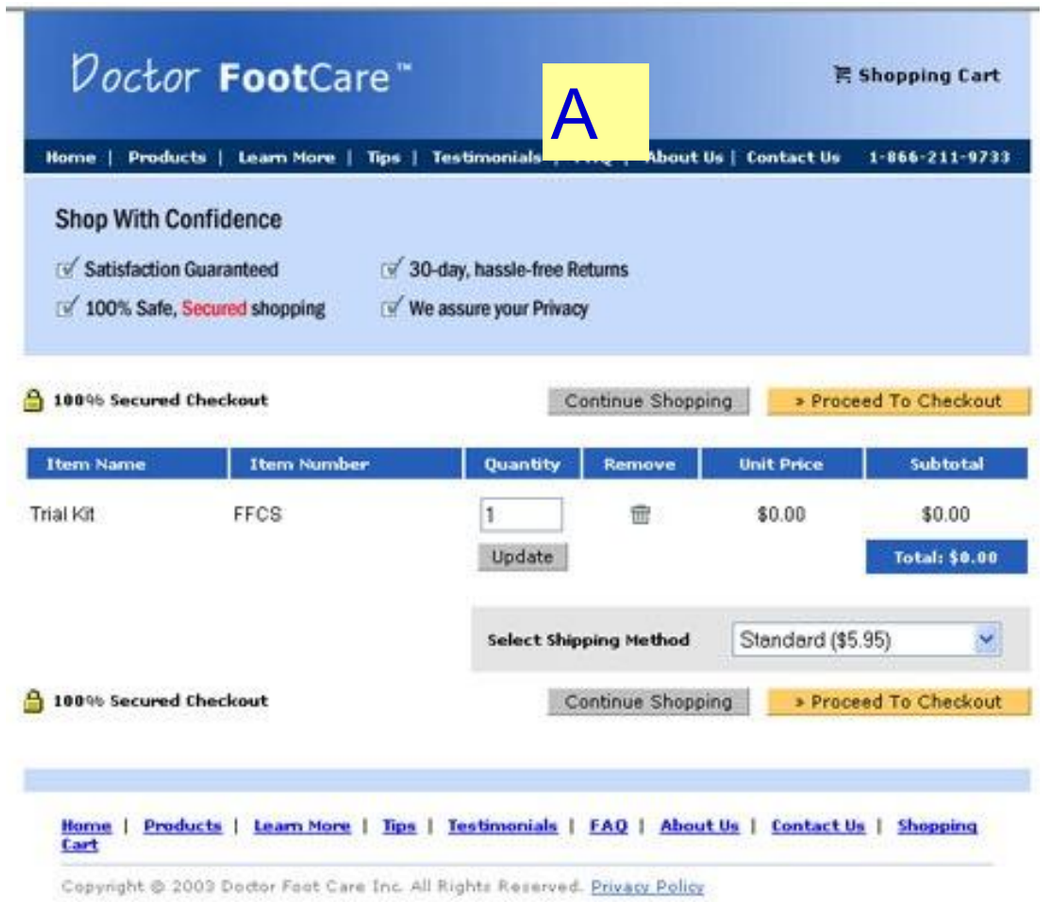
- Stop the project

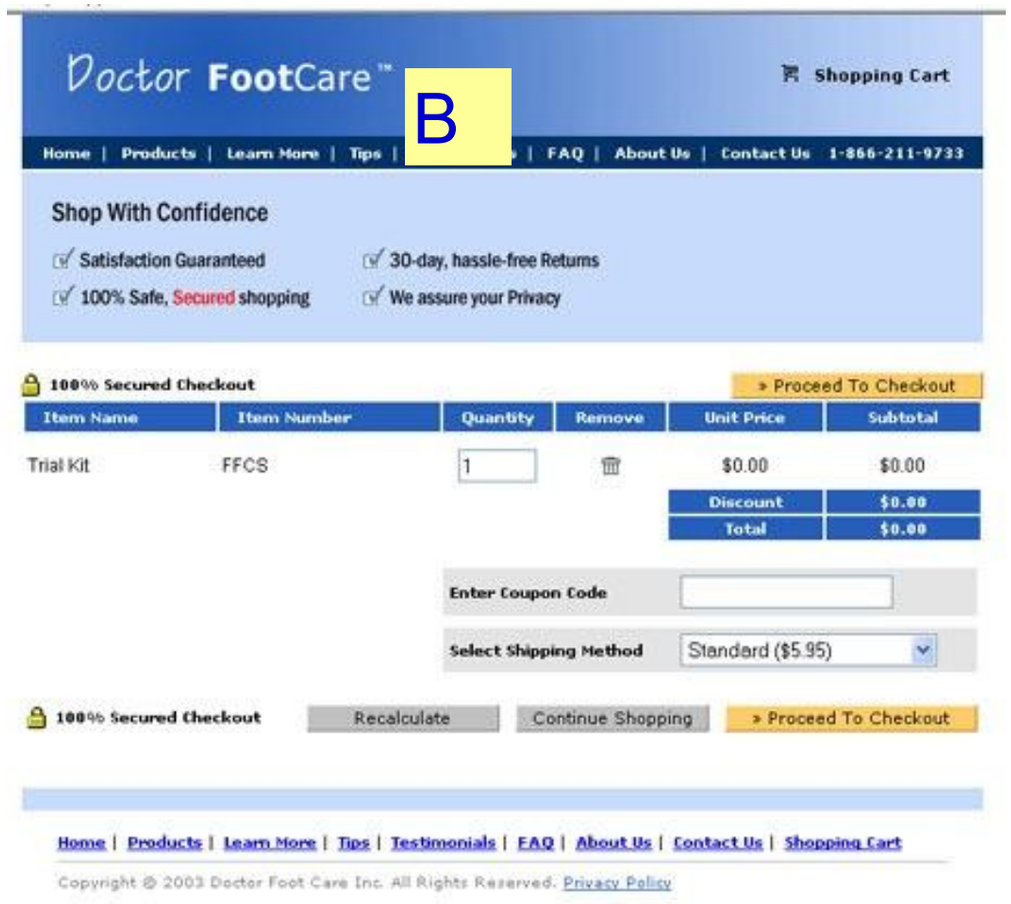
Simple experiment run:

- Wildly successful

Checkout Page

Conversion rate is percentage of visits that include purchase

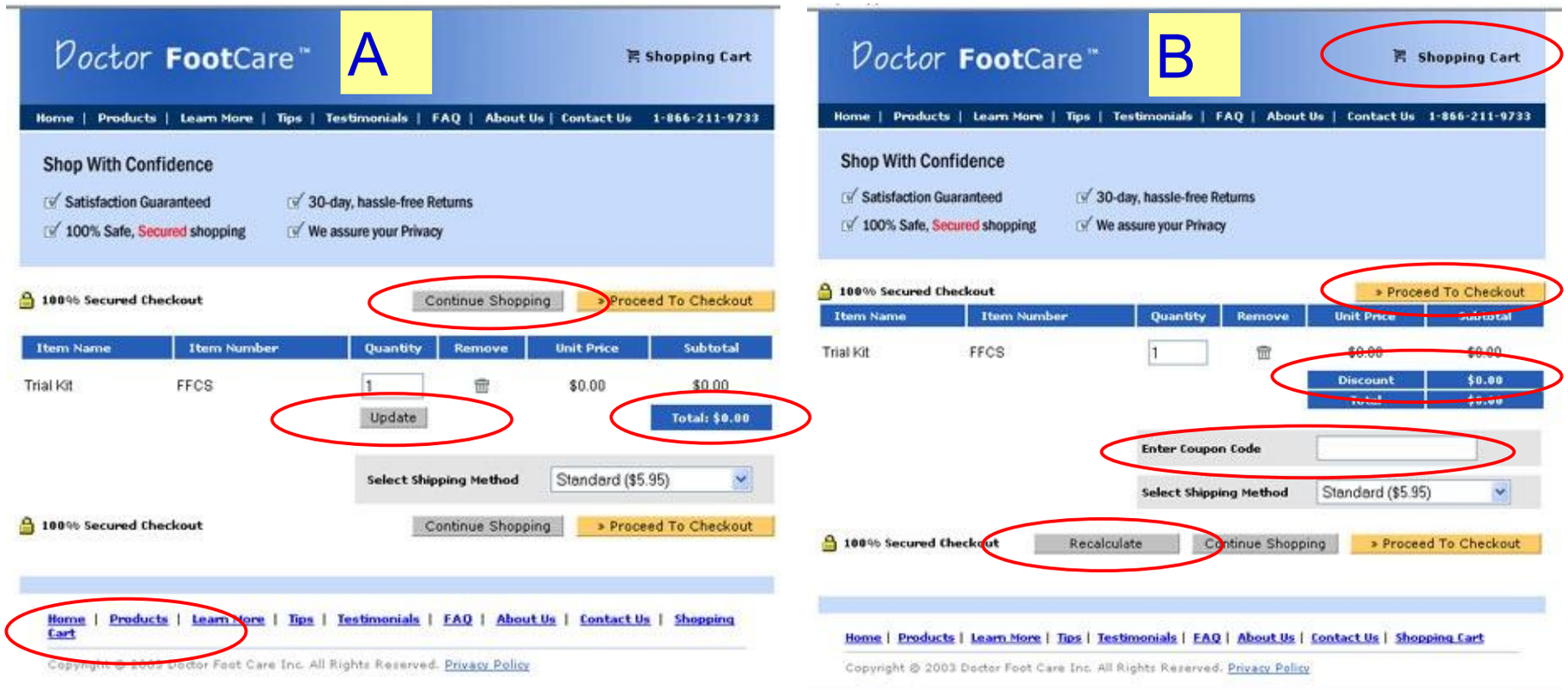




Which version has a higher conversion rate?

Checkout Page

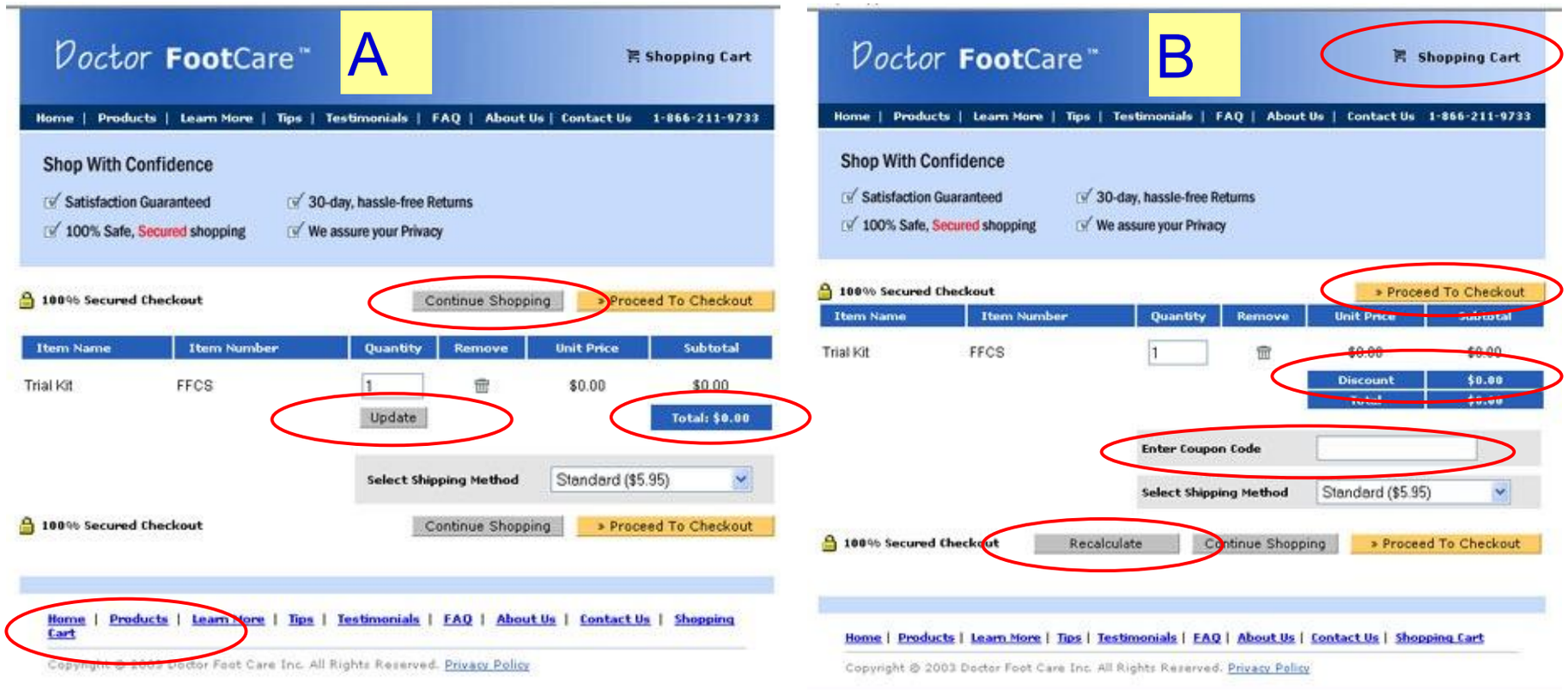
Conversion rate is percentage of visits that include purchase



Which version has a higher conversion rate?

Checkout Page

Conversion rate is percentage of visits that include purchase



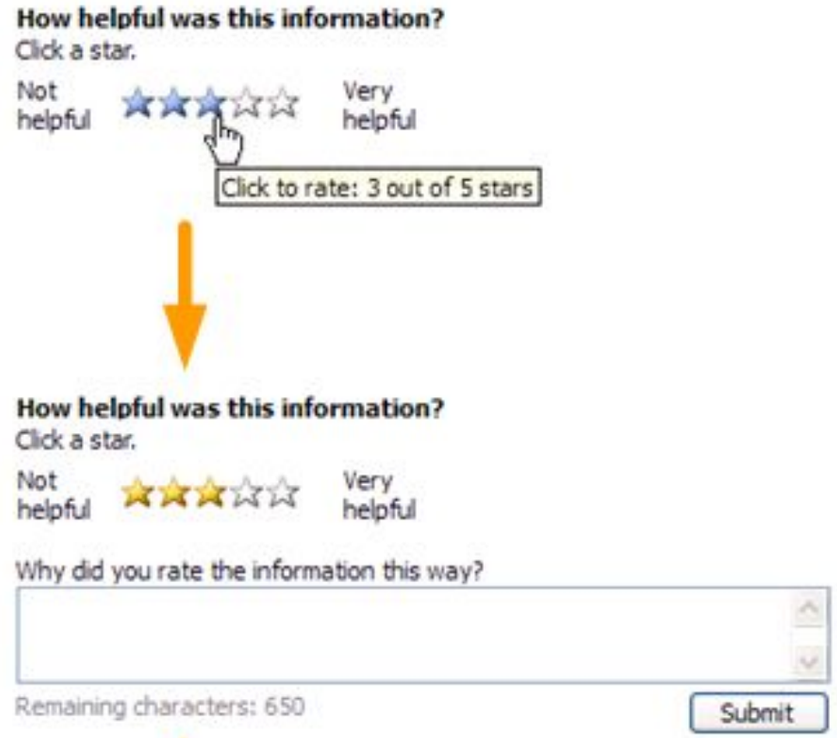
Which version has a higher conversion rate?

Coupon Code decreases by factor of 10

Office Online Feedback

A

B

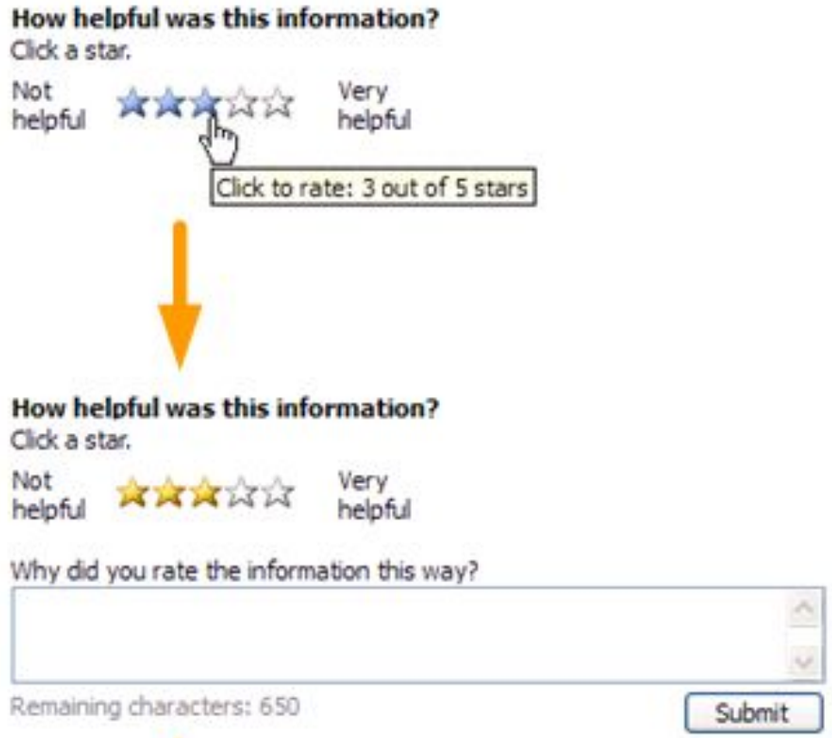


Which one has a higher response rate? By how much?

Office Online Feedback

A

B

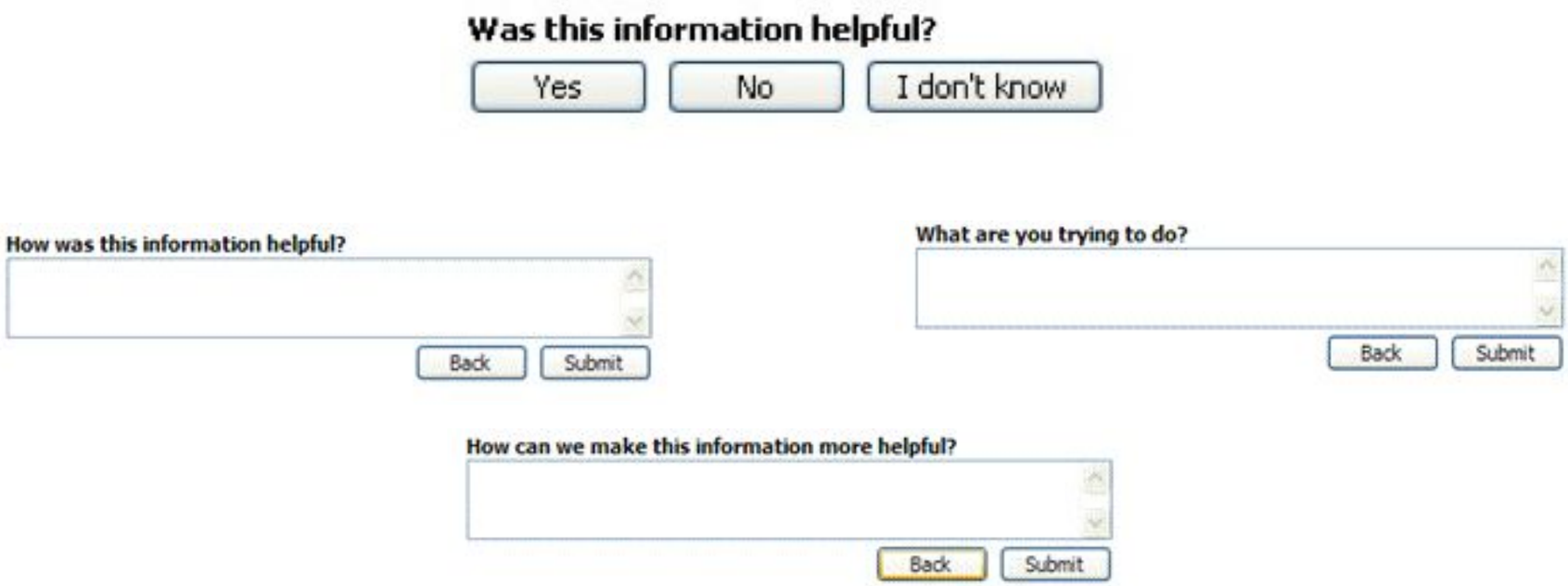


Which one has a higher response rate? By how much?
B gets more than double...

Another Feedback Variant

Call this variant C. Like B, also two stage.

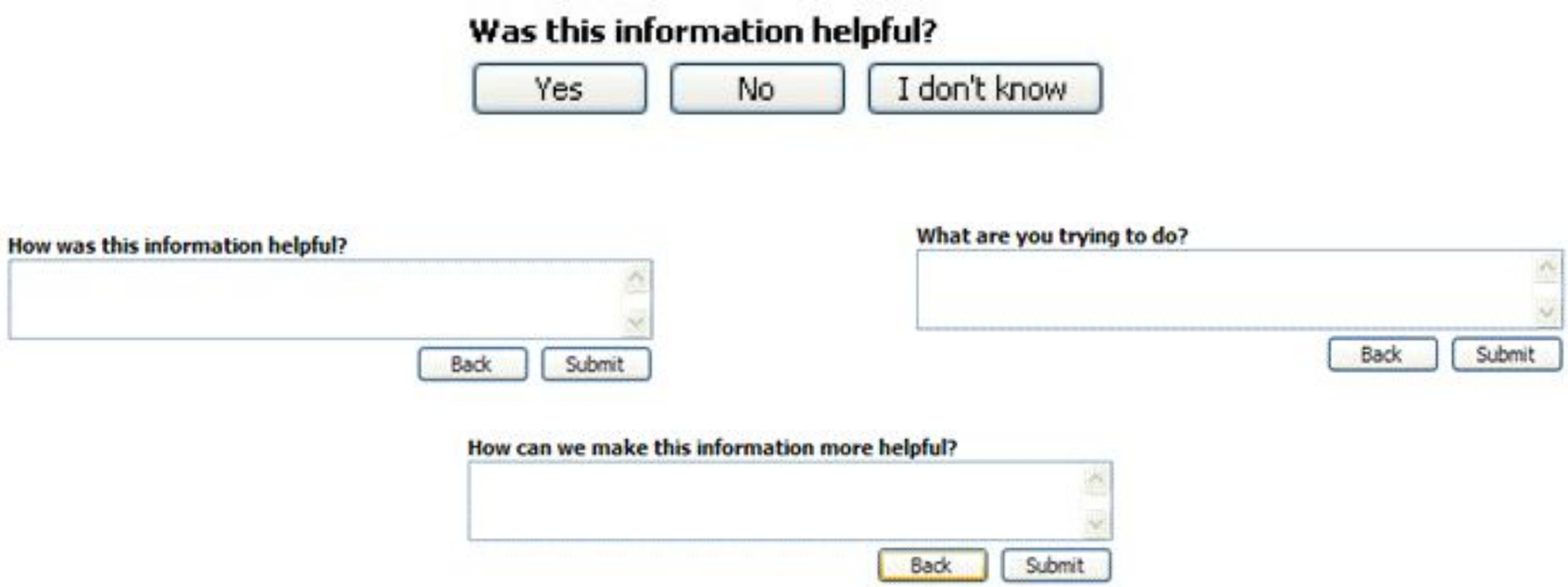
Which one has a higher response rate, B or C?



Another Feedback Variant

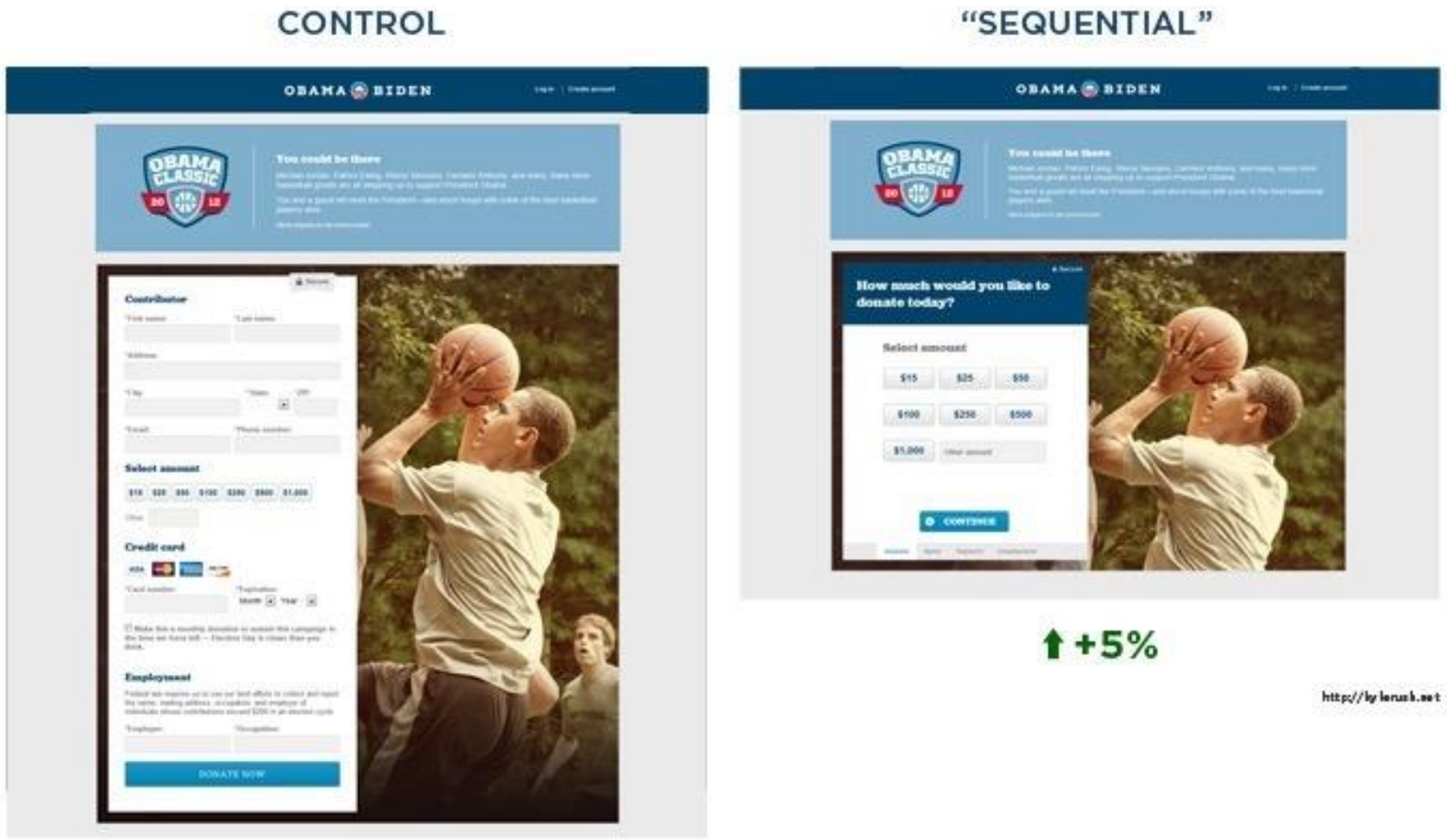
Call this variant C. Like B, also two stage.

Which one has a higher response rate, B or C?



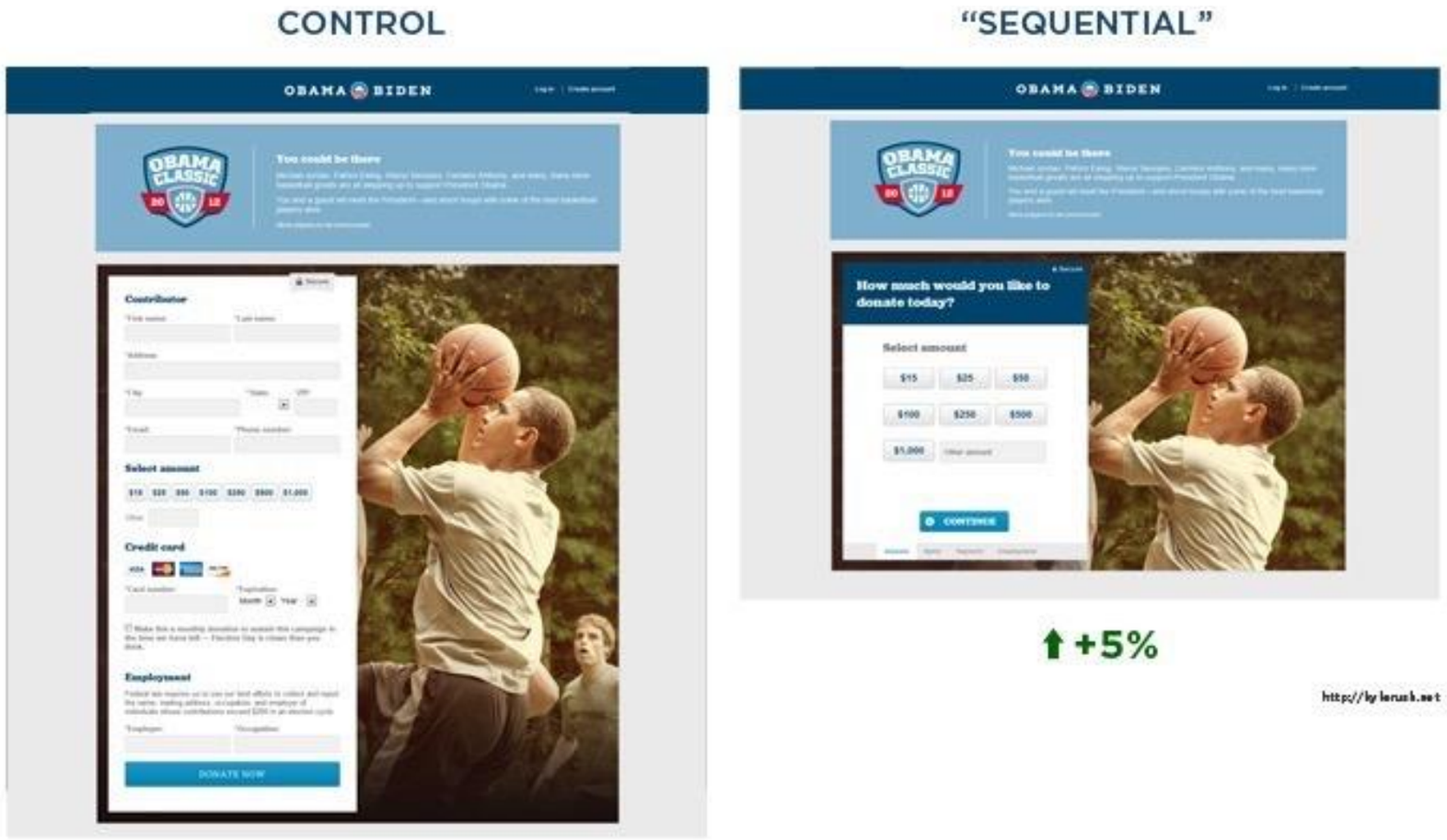
C outperforms B by a factor of 3.5

Obama Campaign



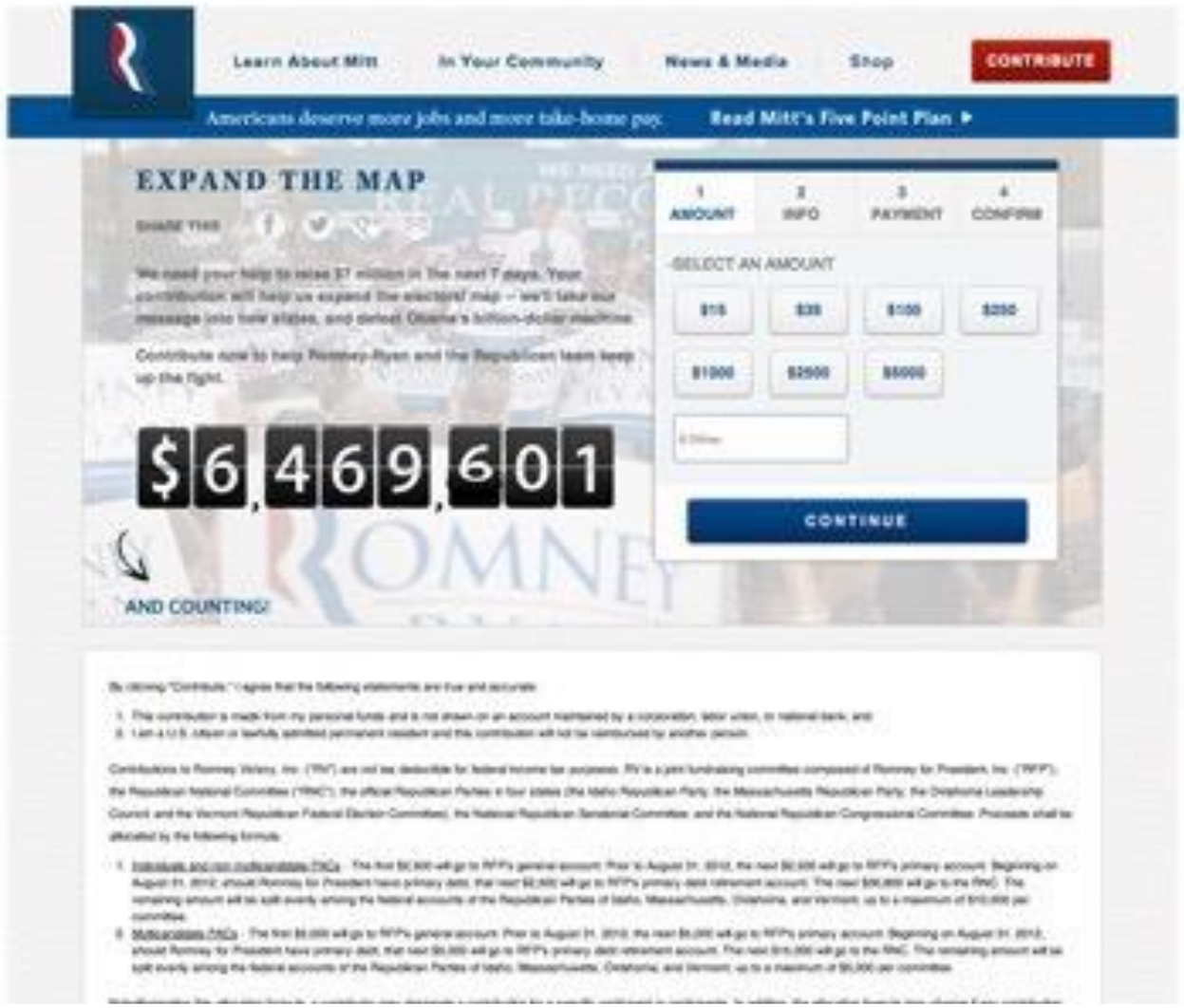
http://o04151985.kylerush.net/blog/optimization-at-the-obama-campaign-ab-testing/

Obama Campaign



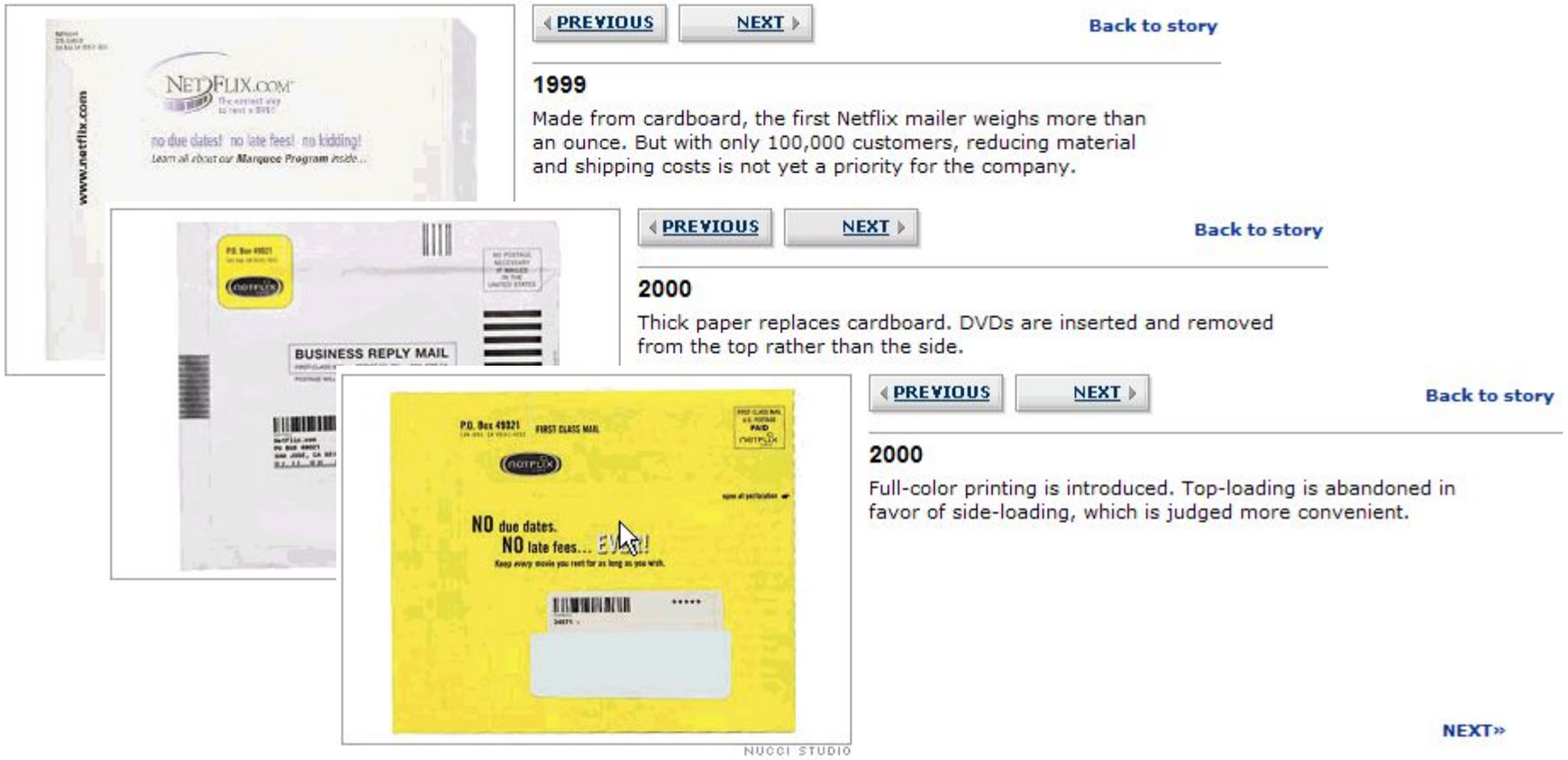
http://o04151985.kylerush.net/blog/optimization-at-the-obama-campaign-ab-testing/

Romney Campaign



http://o04151985.kylerush.net/blog/optimization-at-the-obama-campaign-ab-testing/

Data-Driven Methods Not Just Online



PREVIOUS NEXT Back to story

1999

Made from cardboard, the first Netflix mailer weighs more than an ounce. But with only 100,000 customers, reducing material and shipping costs is not yet a priority for the company.

PREVIOUS NEXT Back to story

2000

Thick paper replaces cardboard. DVDs are inserted and removed from the top rather than the side.

PREVIOUS NEXT Back to story

2000

Full-color printing is introduced. Top-loading is abandoned in favor of side-loading, which is judged more convenient.

NEXT»

Data-Driven Methods Not Just Online

PREVIOUS NEXT Back to story

2000

Customers are asked to peel off a sticker to reveal Netflix's return address. The design is eventually deemed too complex.

PREVIOUS NEXT Back to story

2000

Made from plastic instead of paper, this mailer is cheaper, but it sometimes inflates when transported on airplanes.

PREVIOUS NEXT Back to story

2001

An airhole (the black dot on the left side of the mailer) is added to prevent the package from inflating.

PREVIOUS NEXT Back to story

2001

Netflix returns to paper because it's easier to recycle. Foam padding is added to reduce breakage.

NUCCI STUDIO

NEXT»

Data-Driven Methods Not Just Online

Limitations

Drives hill-climbing, but not overall design

A design may be better, but is it good?

Impossible for new designs to compete

Can be difficult to scale to many features

Ask me something!

Ideation and Evaluation

Case Video:
Phone Button Lay-Out

http://www.youtube.com/watch?v=kCSzjExvbTQ