Too big to fail:

A/B tests when you're betting the bank

Andrea Burbank (@arburbank)

November 4, 2014

Pinterest

#qconsf

March 19, 2013

holed up for 6 months

replace $\frac{2}{3}$ of the code

new features can’t ship

desperately want to release

many things have changed

How did we get here?

What have we done so far?

What should we do **right now**?

<how we got here>

Rewrite the website

MAINTAINABLE

ENCAPSULATED

TESTABLE

Rewrite the website

EXTENSIBLE

COHESIVE

MODULAR

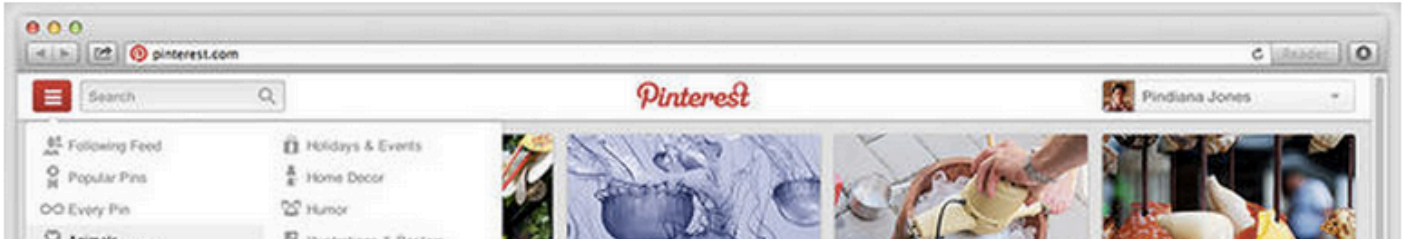
Pinterest tests new look with bigger pins, restructured navigation

The addicting social-networking site is mixing things up a bit to make its pins even more gripping.

by Jennifer Van Grove @jbruin / January 28, 2013 3:02 PM PST

Pinterest is readying a test that will alter the inspirational site with a tweaked navigation, bigger pins, and an overall faster site experience, the company announced Monday.

The **experimental Pinterest look** will roll out to a small group of people as the startup attempts to help its users get more information about pins they care about, a spokesperson told CNET. The company plans to release the design more widely once it works out the bugs.



http://www.cnet.com/news/pinterest-tests-new-look-with-bigger-pins-restructured-navigation/

moved all the buttons

rounded all the corners

replaced the categories menu

made the pins bigger

removed friends to follow

added related pins to closeup

got rid of the “add a pin” menu

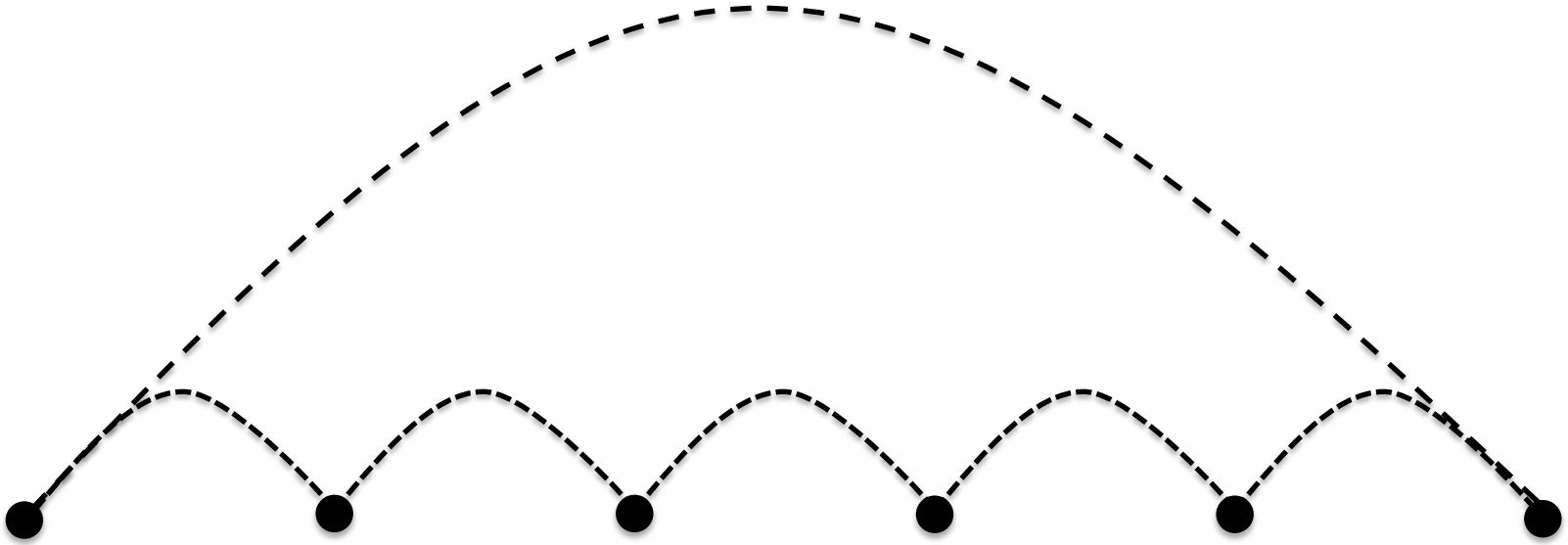
removed second-order attribution

renamed the Repin button

uncluttered the pin closeup

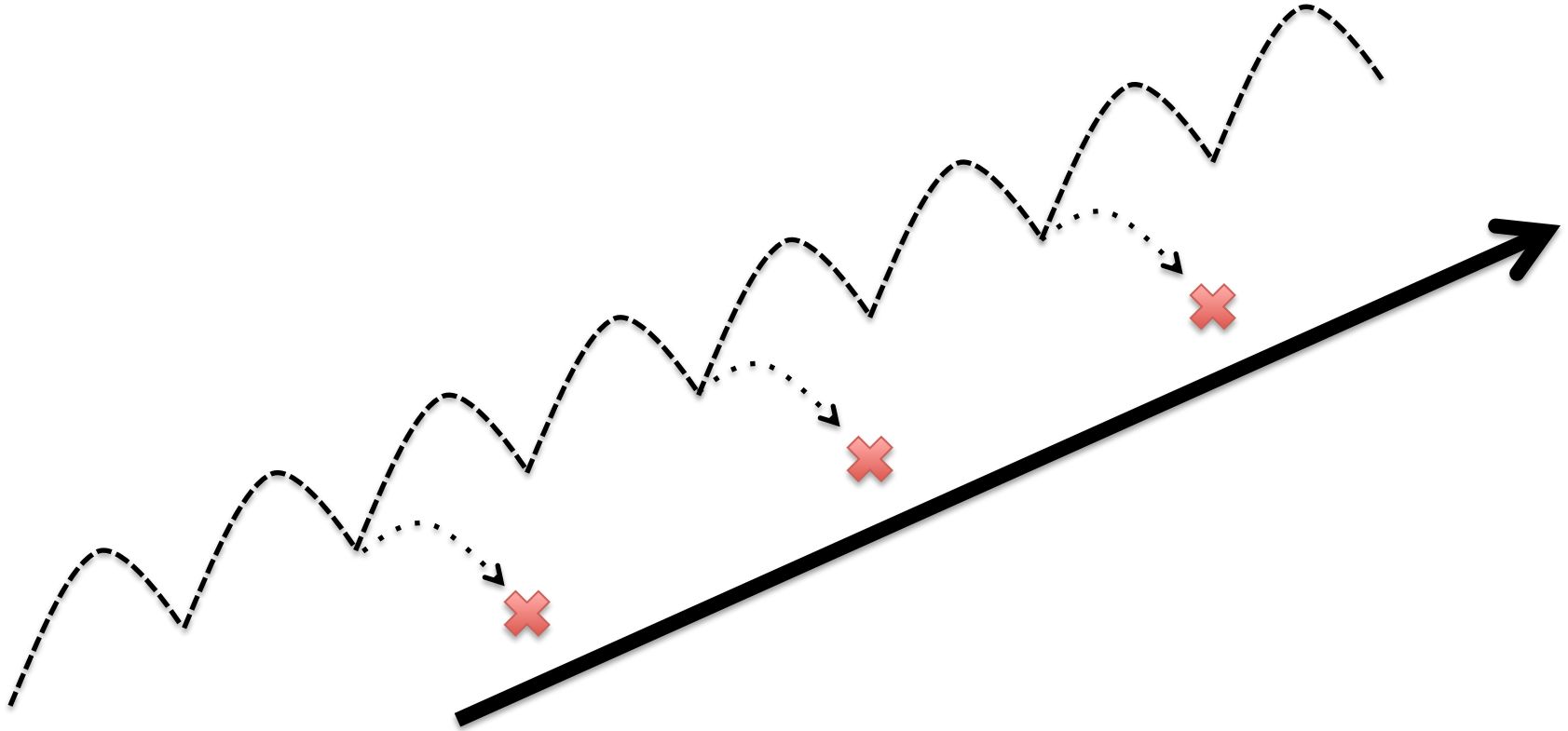
and much more

Why so much change at once?



“sometimes, don’t shoot yourself in the foot.”

Advantages of incremental change

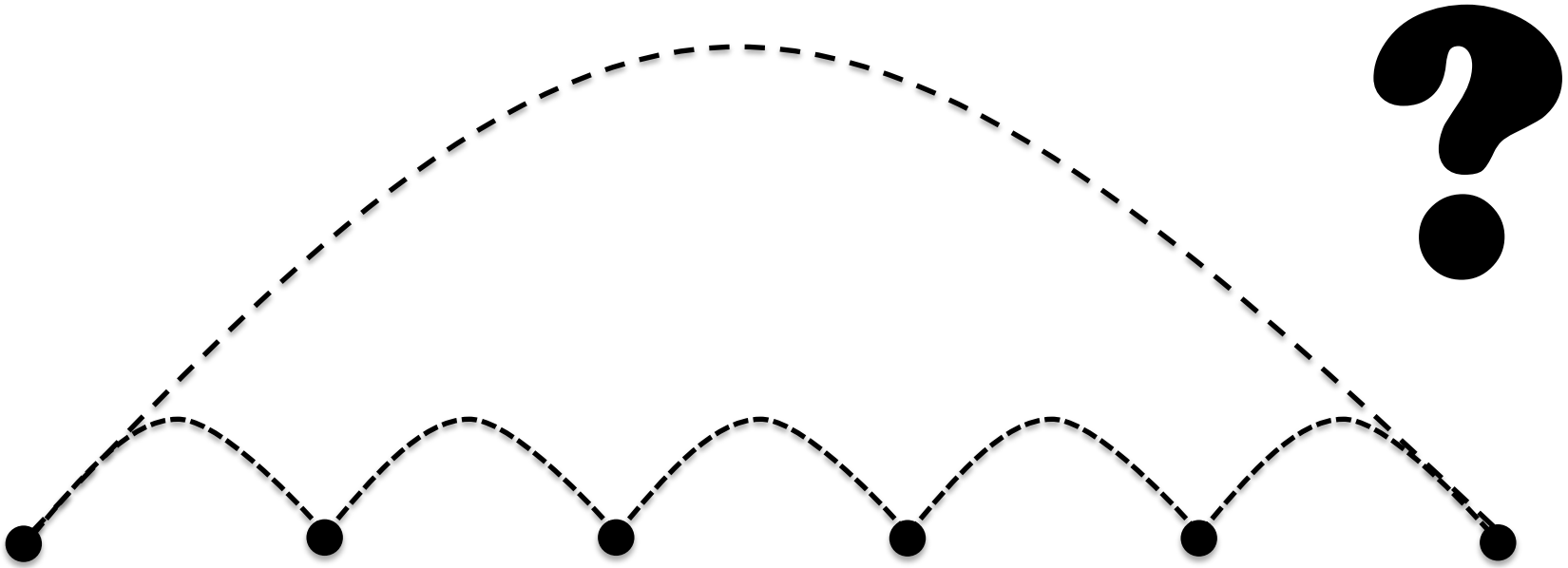


confidence

extensible learnings

focus your resources

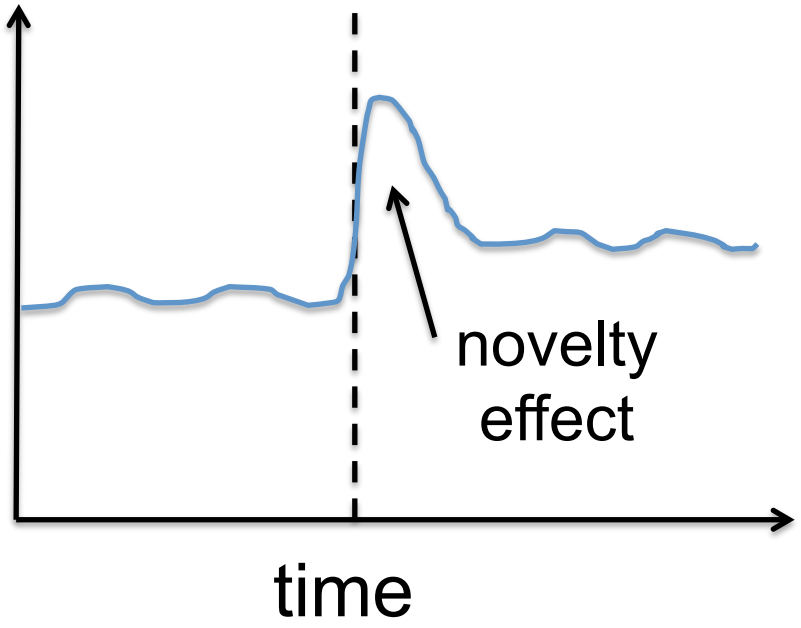
Incremental or all at once?



understand (and rebut) arguments against incremental change

avoid our mistakes

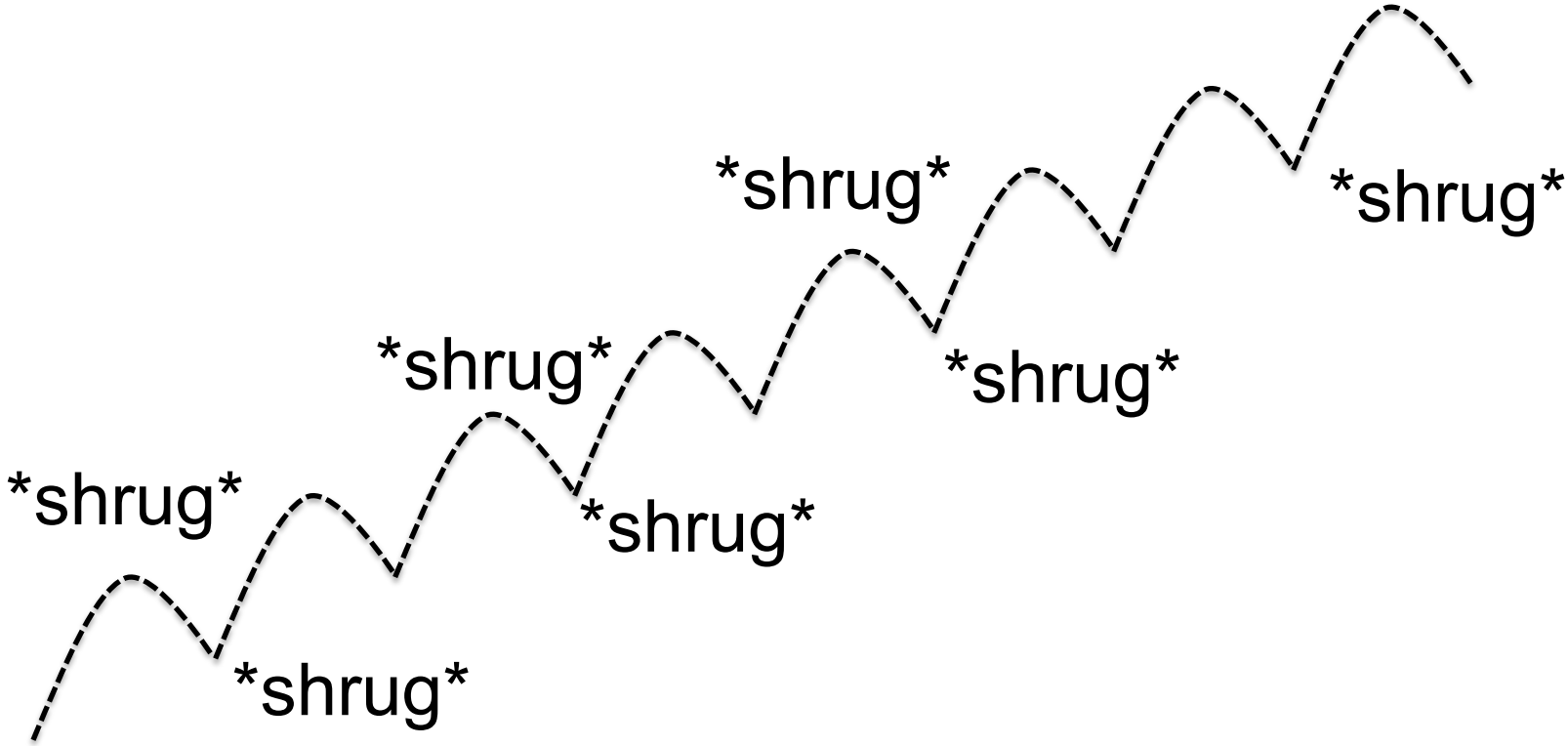
Cons of small changes #1: **time**



rebuttal #1: **time**

big changes will be slower than you anticipated

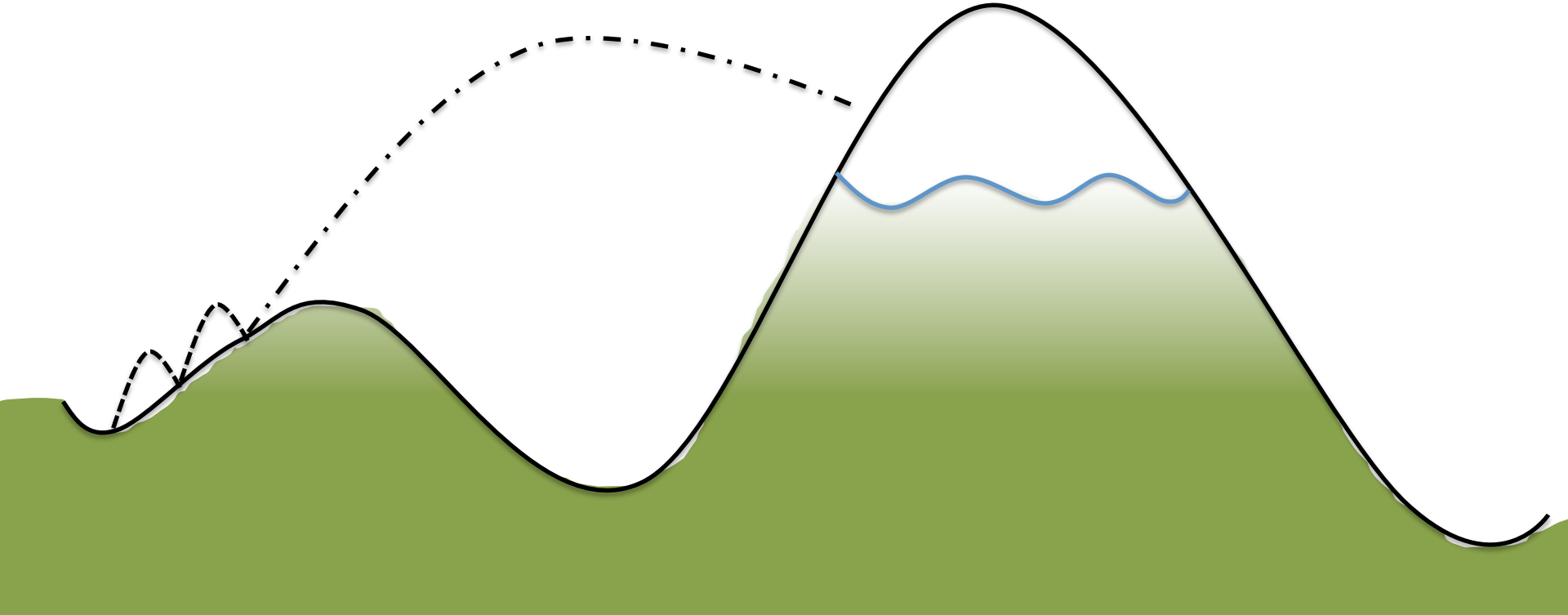
Cons of small changes #2: **psychology**



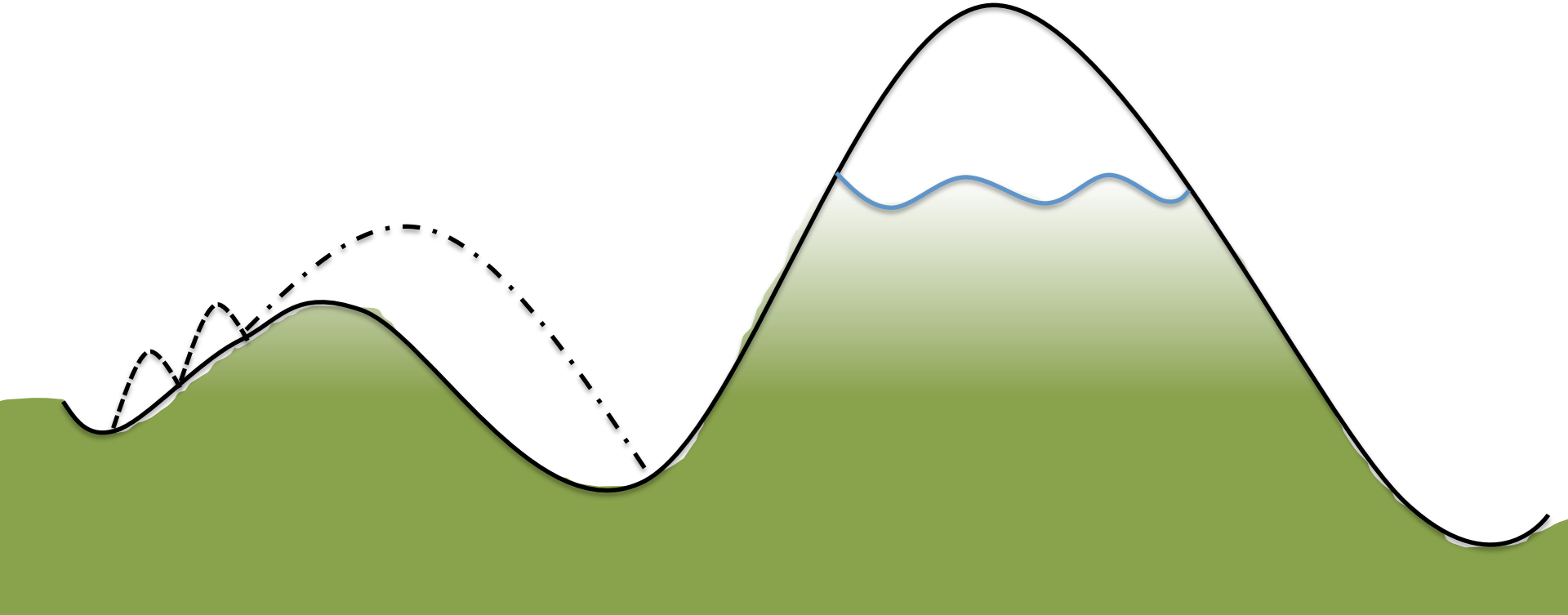
rebuttal #2: **psychology**

it’s not “incremental change” –
it’s **faster learning**

Cons of small changes #3: local maxima

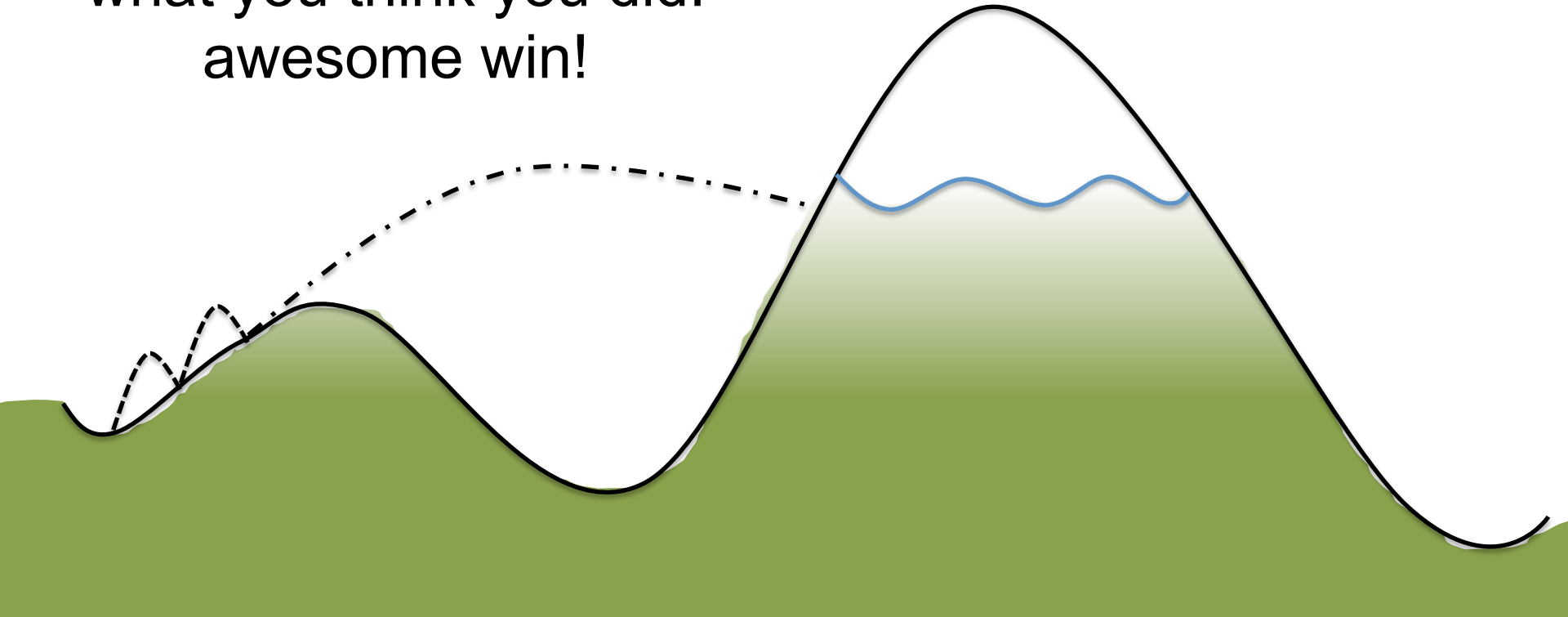


Cons of small changes #3: local maxima

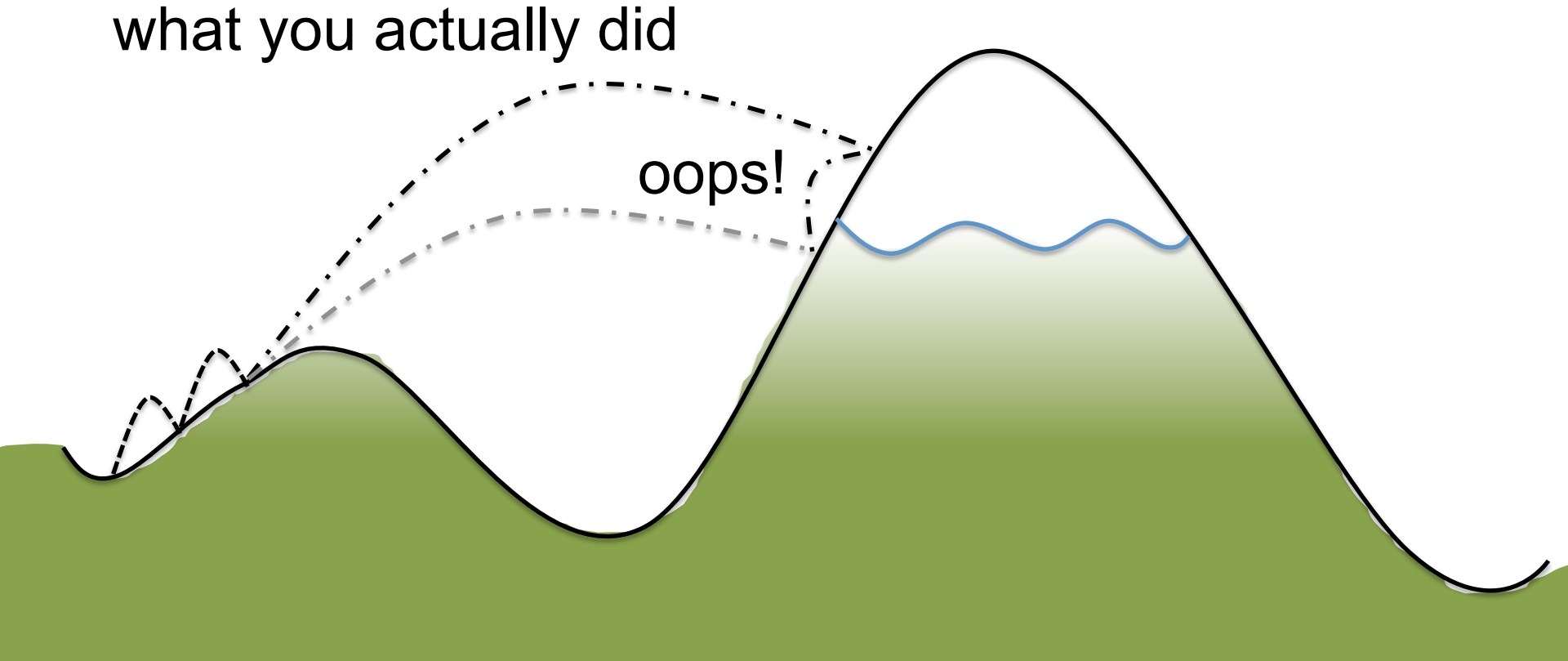


Cons of small changes #3: local maxima

what you think you did:
awesome win!



Cons of small changes #3: local maxima



rebuttal #3: **local maxima**

mountains and valleys
are neighbors

Incremental vs. all at once?

time

avoid false optimism: big changes take time too

psychology

help your organization embrace fast learning

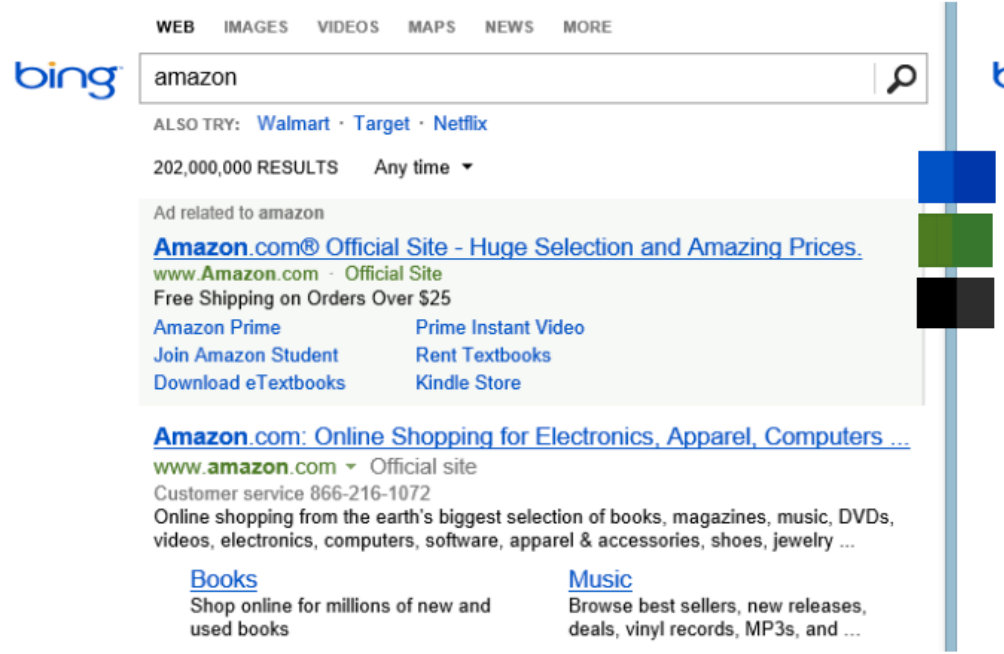
local maxima

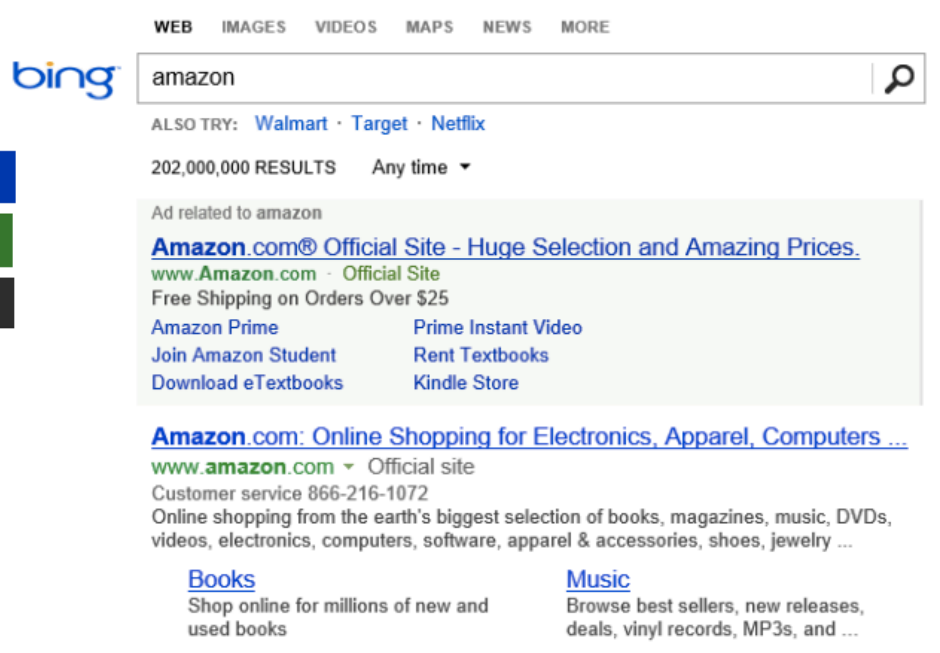
yes! but avoid the valley, and watch your step

a big leap

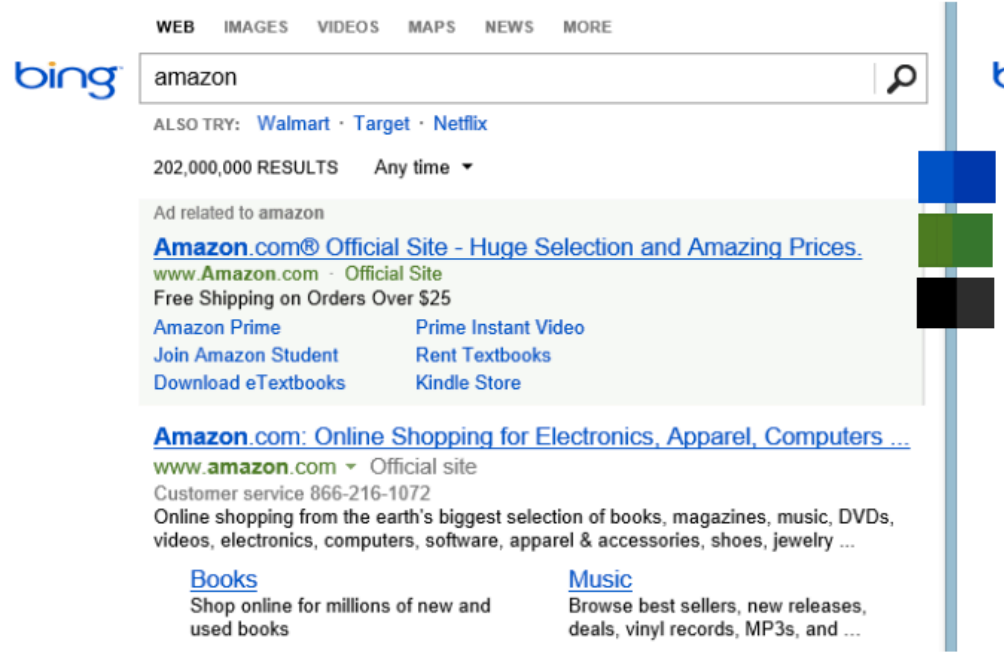
AT&T
9:42 AM
SMS
Text
9
Calendar
Photos
Camera
Stocks
Maps
Weather
Clock
Calculator
Notes
Settings
Phone
Mail
Safari
iPod

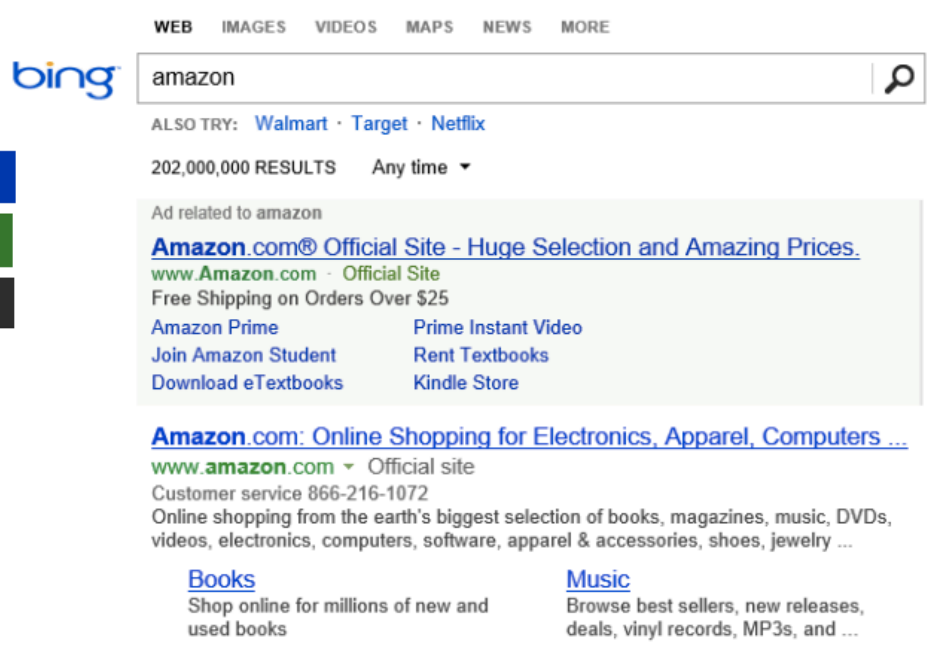
a small change



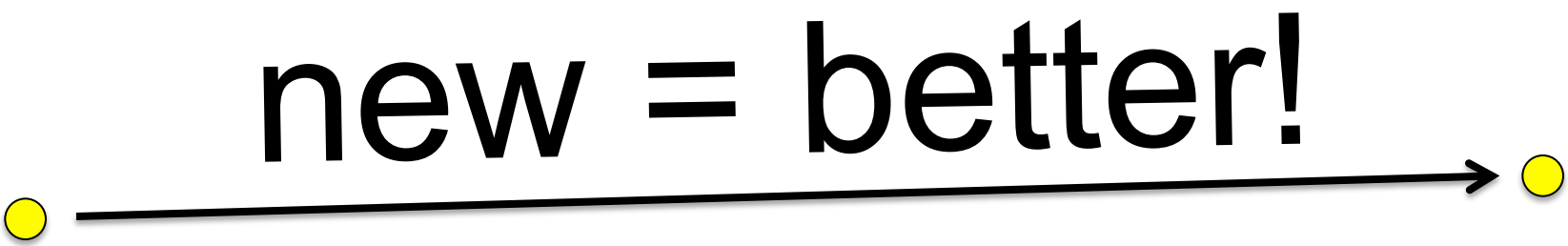


a $10M change


WEB IMAGES VIDEOS MAPS NEWS MORE
bing
amazon
ALSO TRY: Walmart · Target · Netflix
202,000,000 RESULTS Any time
Ad related to amazon
Amazon.com® Official Site - Huge Selection and Amazing Prices.
www.Amazon.com · Official Site
Free Shipping on Orders Over $25
Amazon Prime
Prime Instant Video
Join Amazon Student
Rent Textbooks
Download eTextbooks
Kindle Store
Amazon.com: Online Shopping for Electronics, Apparel, Computers ...
www.amazon.com · Official site
Customer service 866-216-1072
Online shopping from the earth's biggest selection of books, magazines, music, DVDs, videos, electronics, computers, software, apparel & accessories, shoes, jewelry ...
Books
Shop online for millions of new and used books
Music
Browse best sellers, new releases, deals, vinyl records, MP3s, and ...

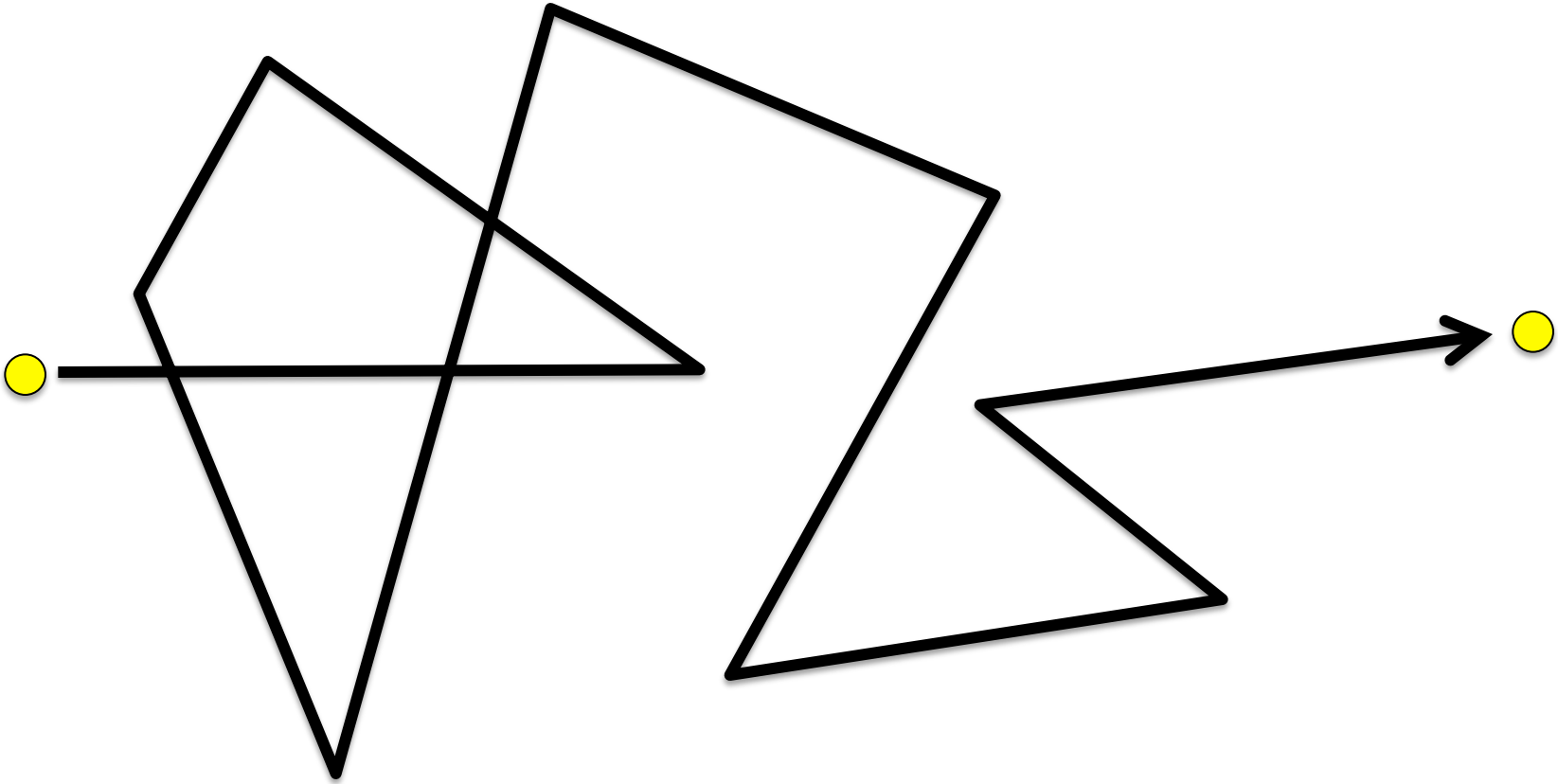

WEB IMAGES VIDEOS MAPS NEWS MORE
bing
amazon
ALSO TRY: Walmart · Target · Netflix
202,000,000 RESULTS Any time
Ad related to amazon
Amazon.com® Official Site - Huge Selection and Amazing Prices.
www.Amazon.com · Official Site
Free Shipping on Orders Over $25
Amazon Prime
Prime Instant Video
Join Amazon Student
Rent Textbooks
Download eTextbooks
Kindle Store
Amazon.com: Online Shopping for Electronics, Apparel, Computers ...
www.amazon.com · Official site
Customer service 866-216-1072
Online shopping from the earth's biggest selection of books, magazines, music, DVDs, videos, electronics, computers, software, apparel & accessories, shoes, jewelry ...
Books
Shop online for millions of new and used books
Music
Browse best sellers, new releases, deals, vinyl records, MP3s, and ...

Making a big change all at once

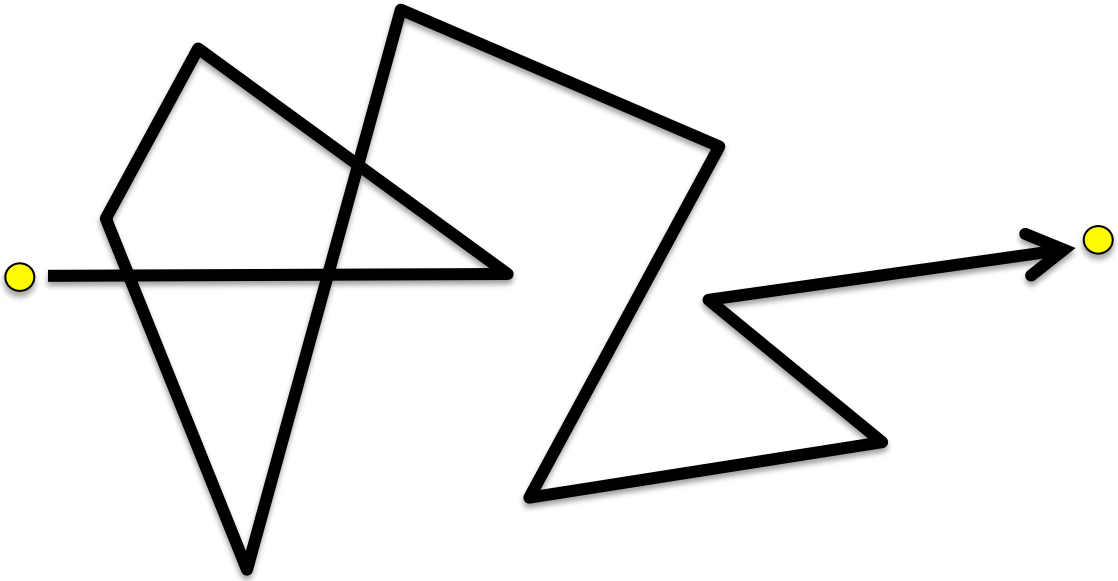


Making a big change all at once

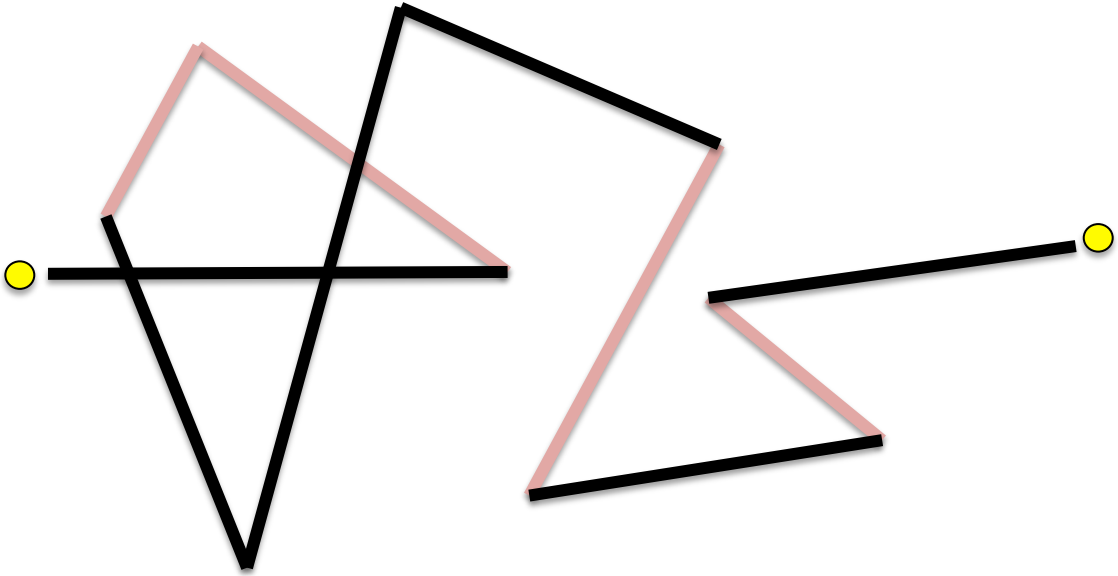
Making a big change all at once



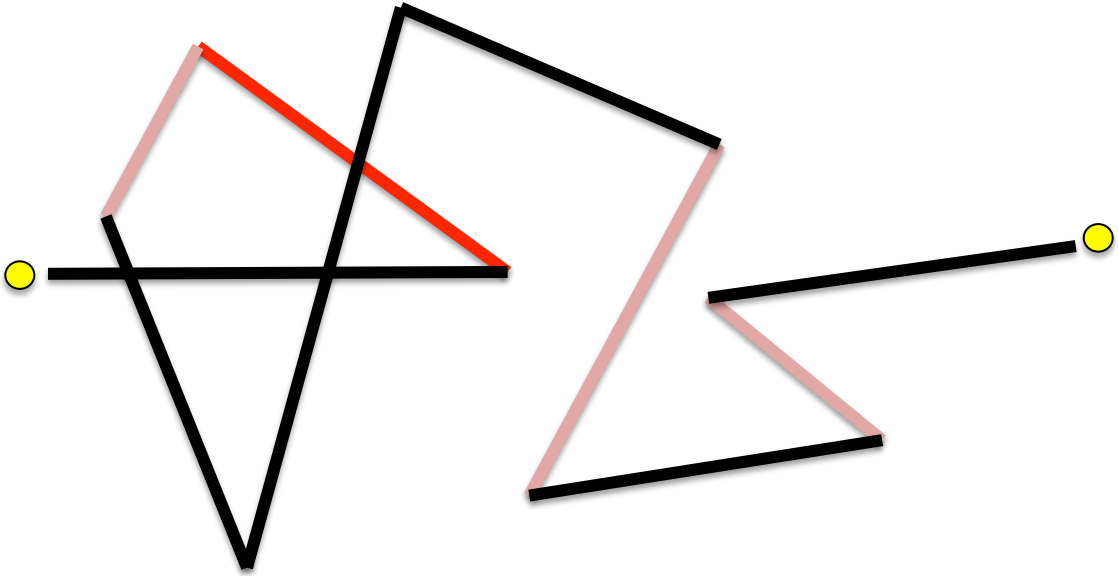
Making a big change all at once



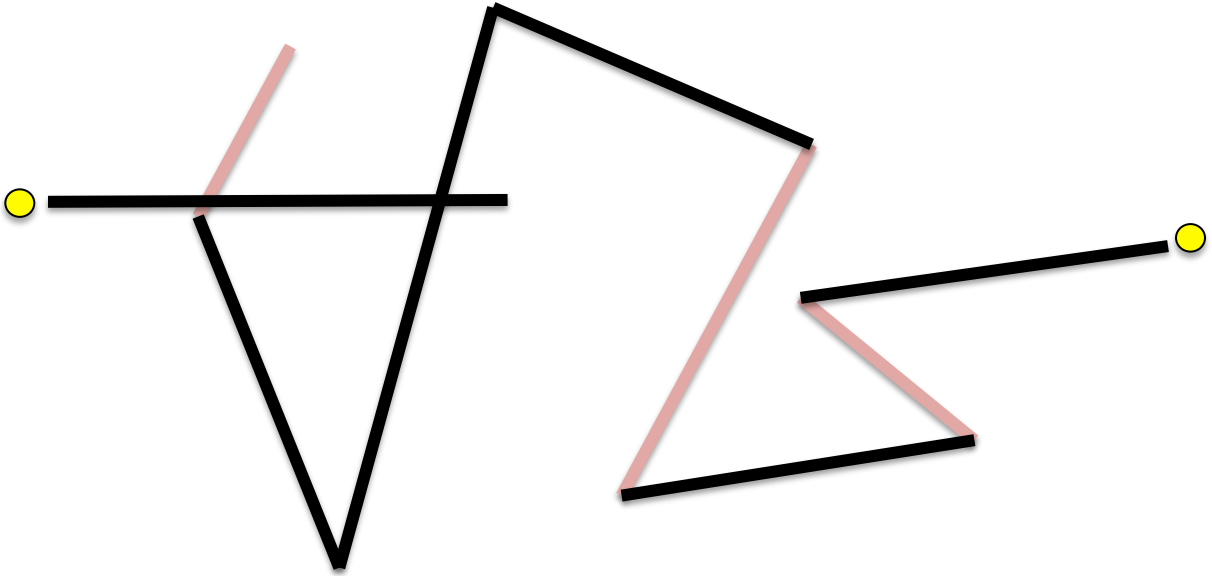
Making a big change all at once



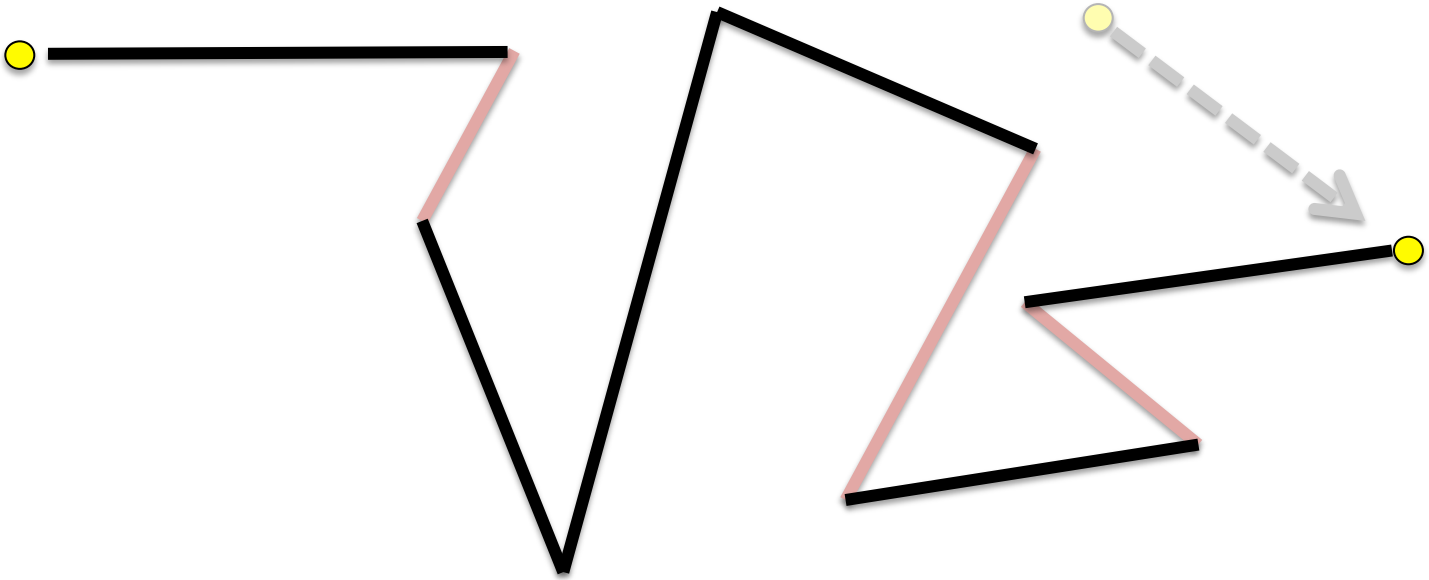
Making a big change all at once



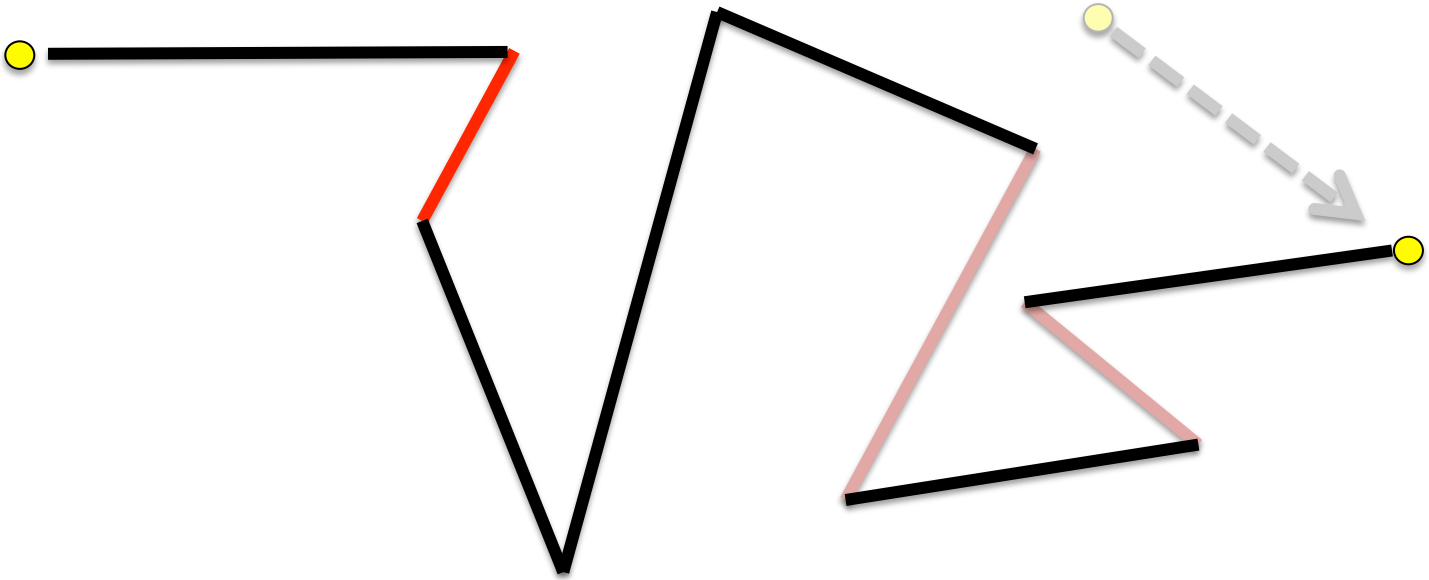
Making a big change all at once



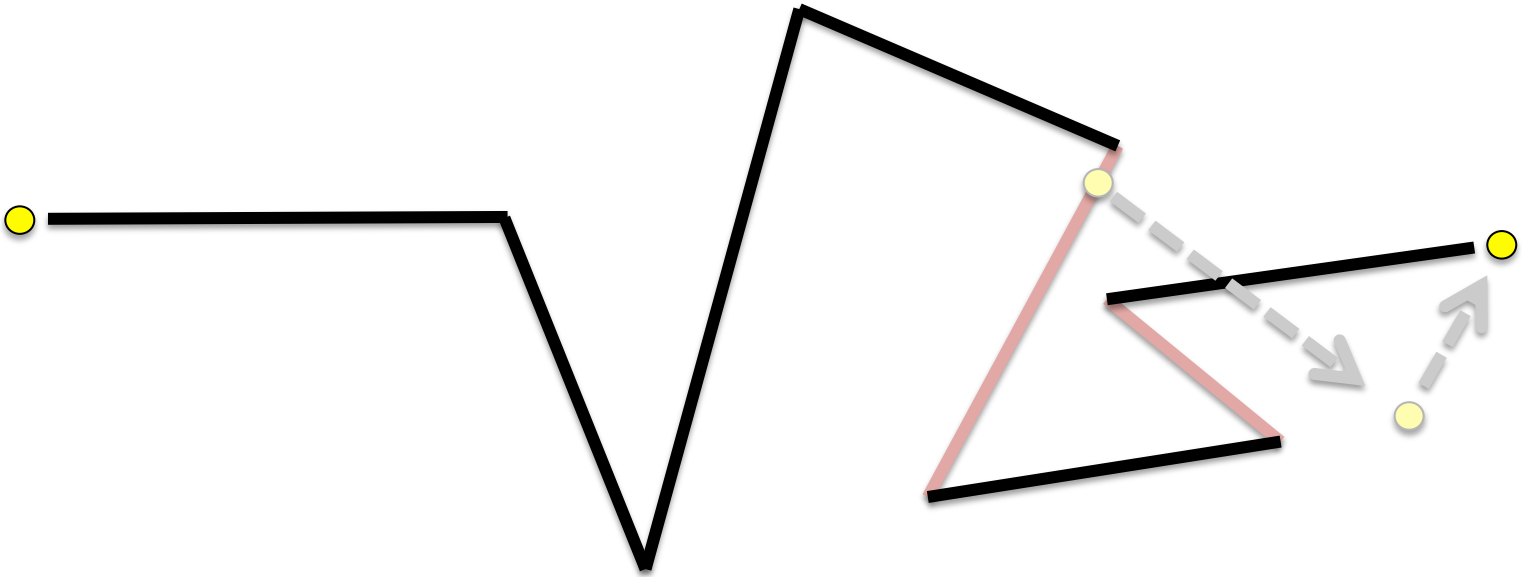
Making a big change all at once



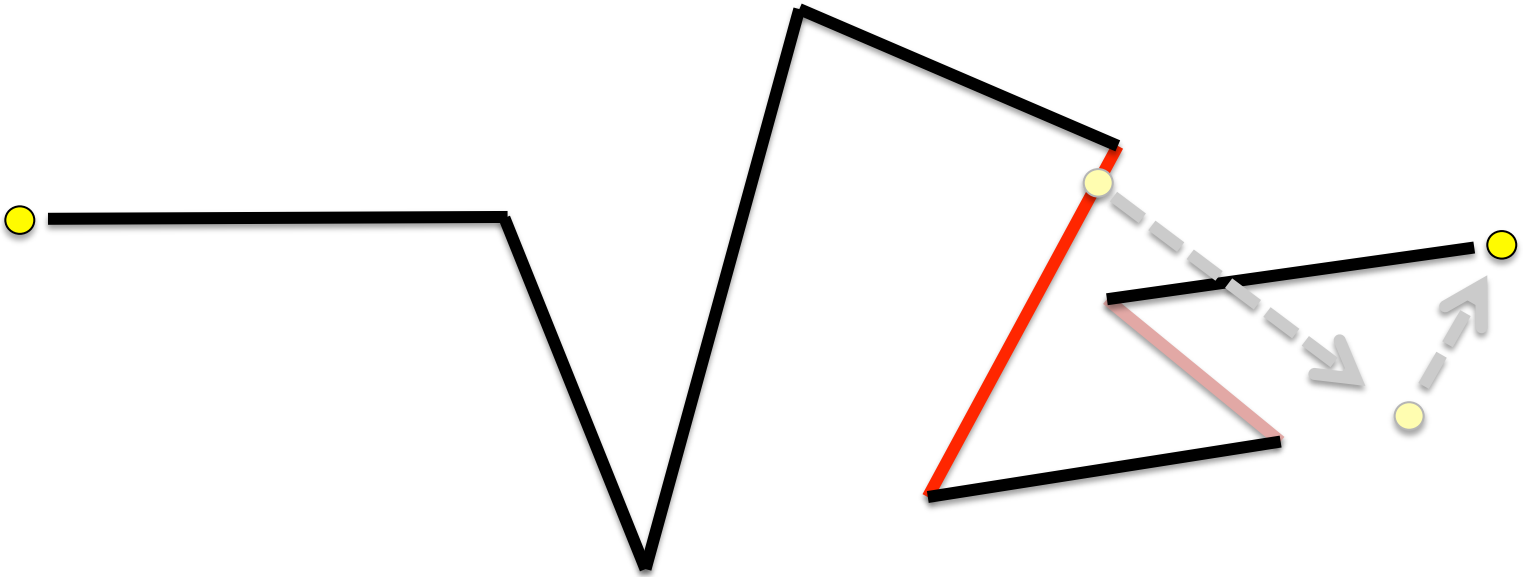
Making a big change all at once



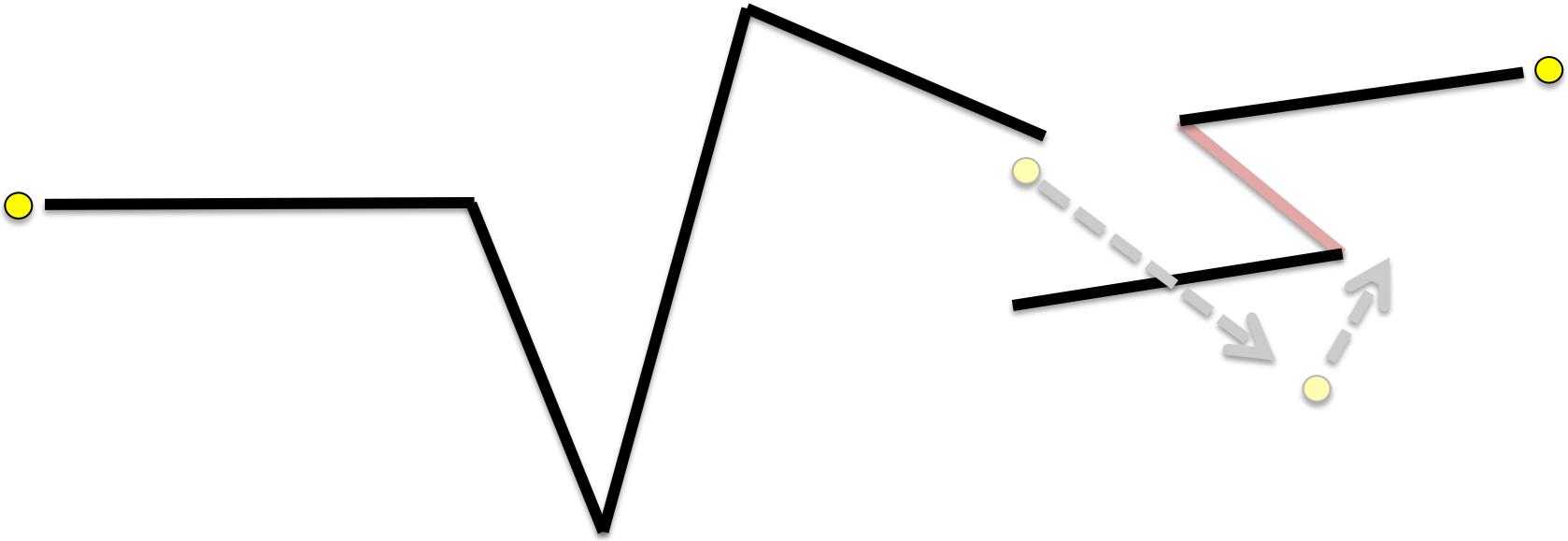
Making a big change all at once



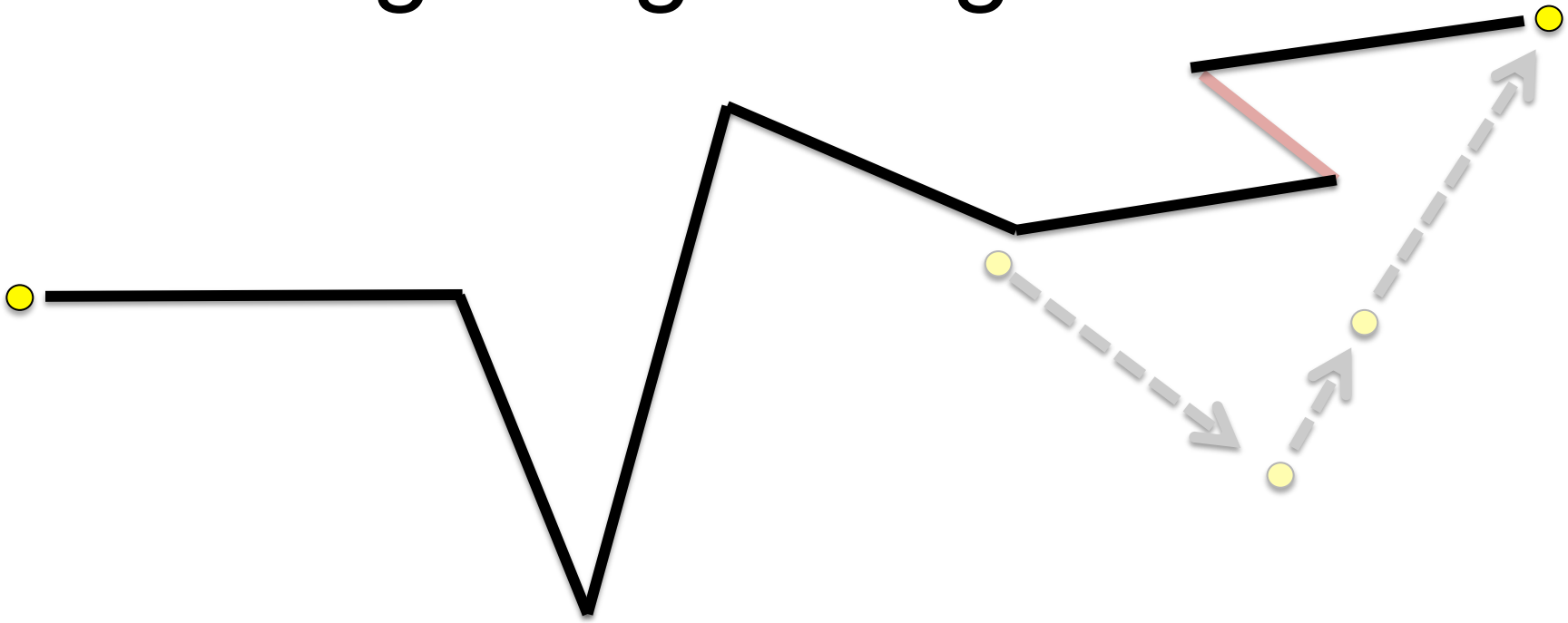
Making a big change all at once



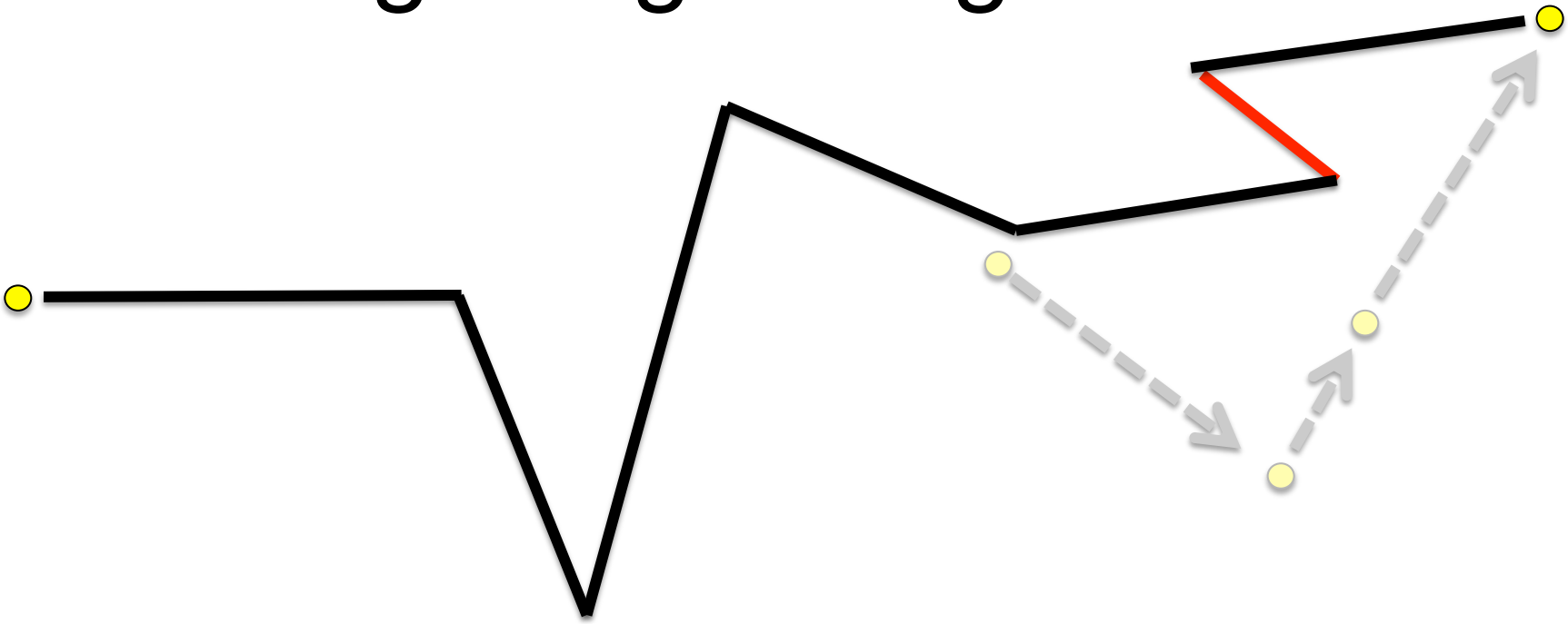
Making a big change all at once



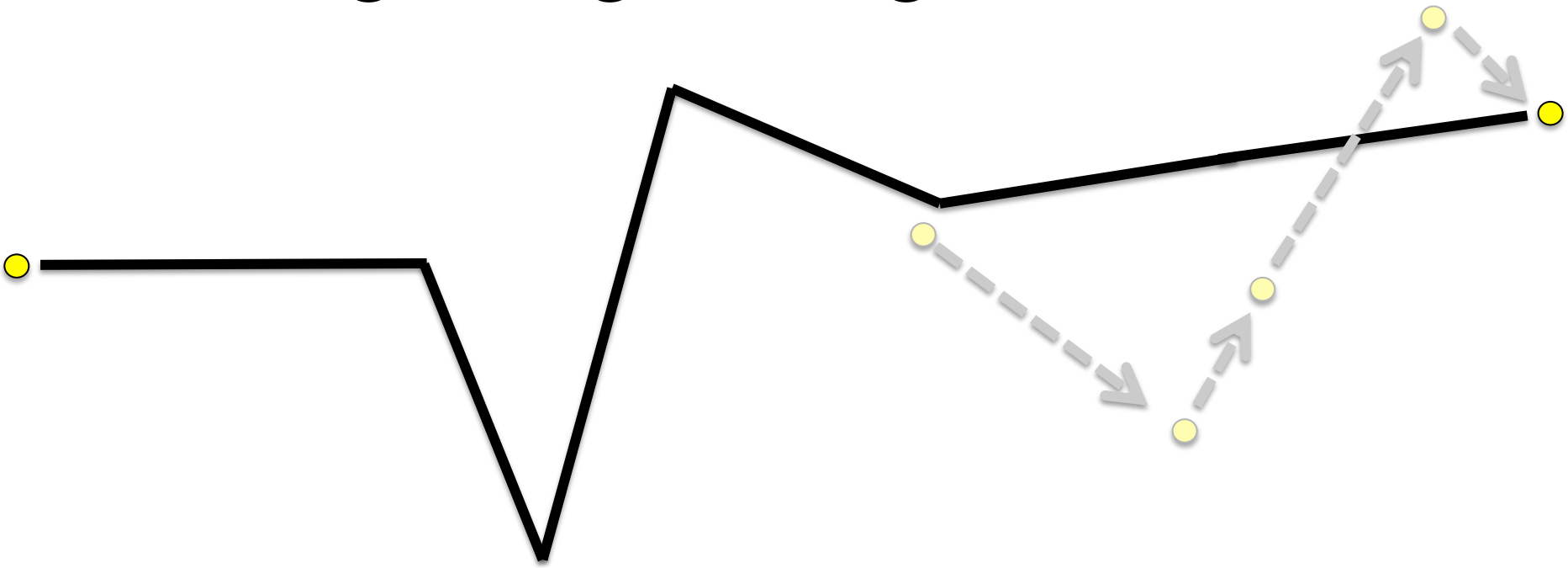
Making a big change all at once



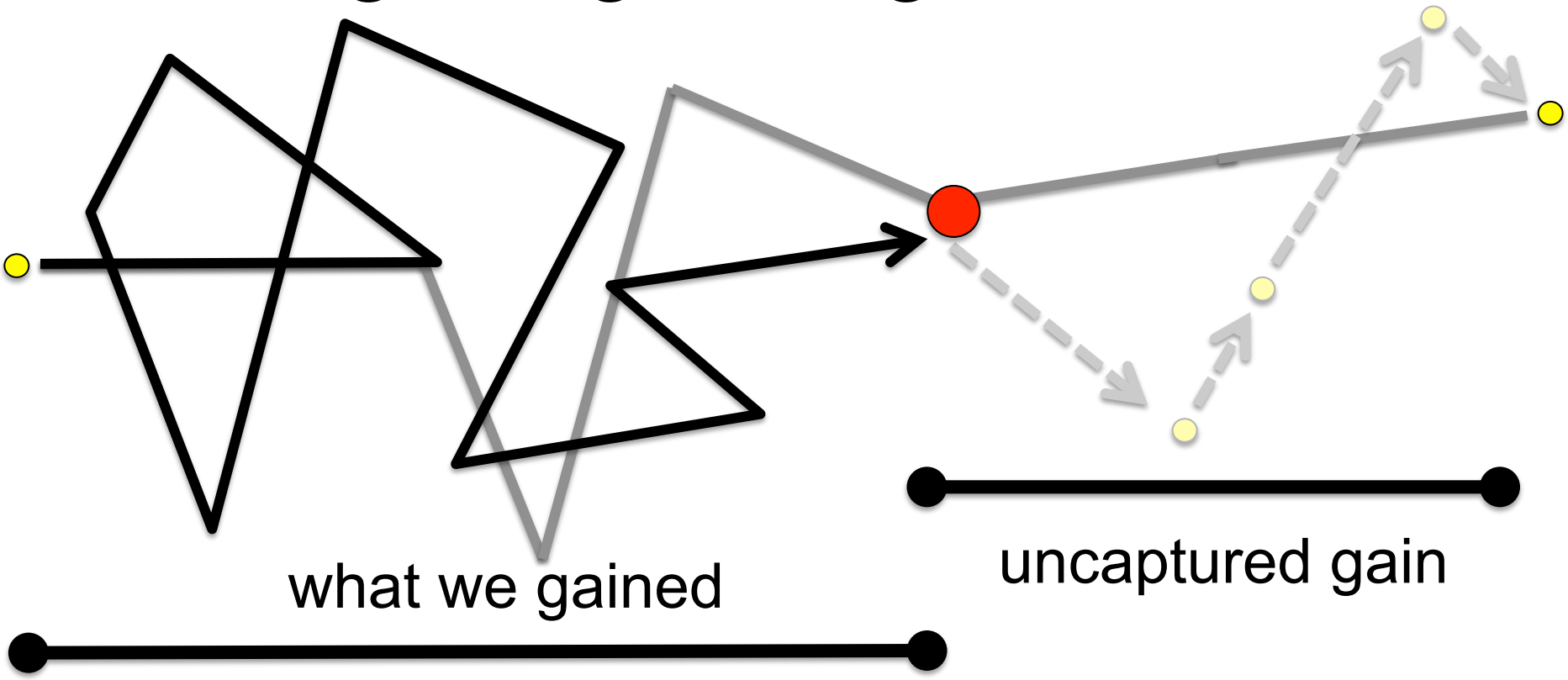
Making a big change all at once



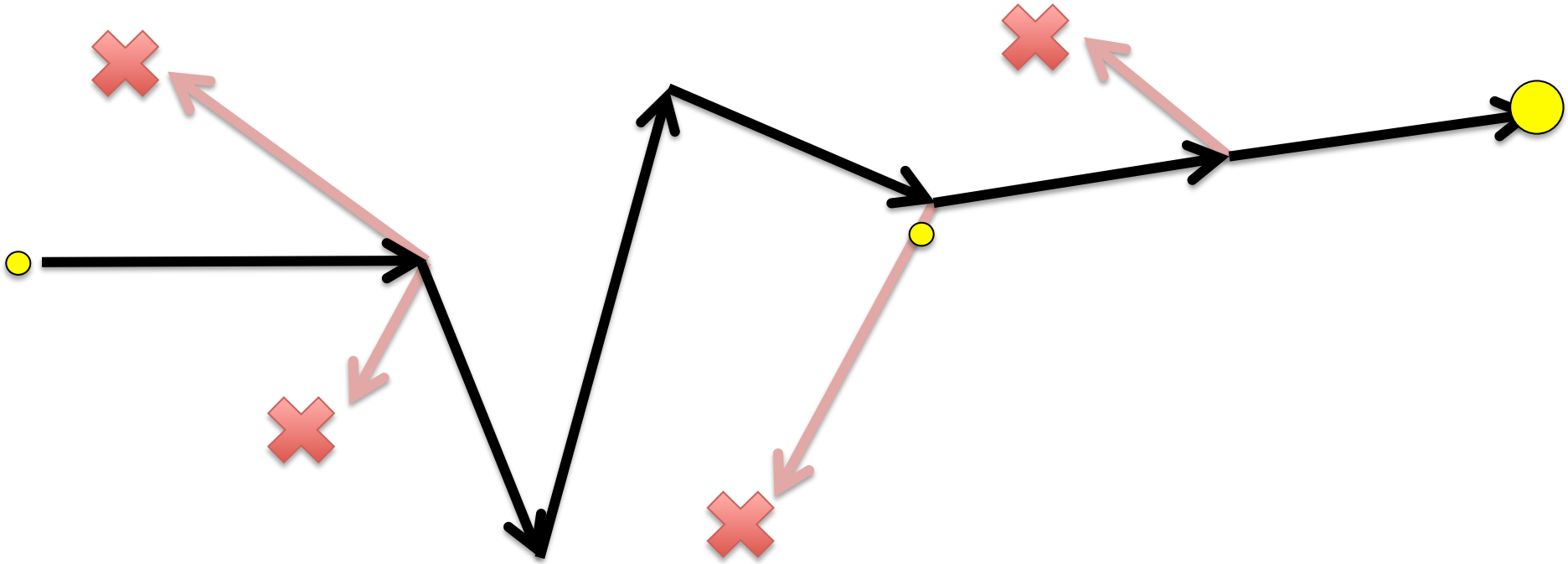
Making a big change all at once



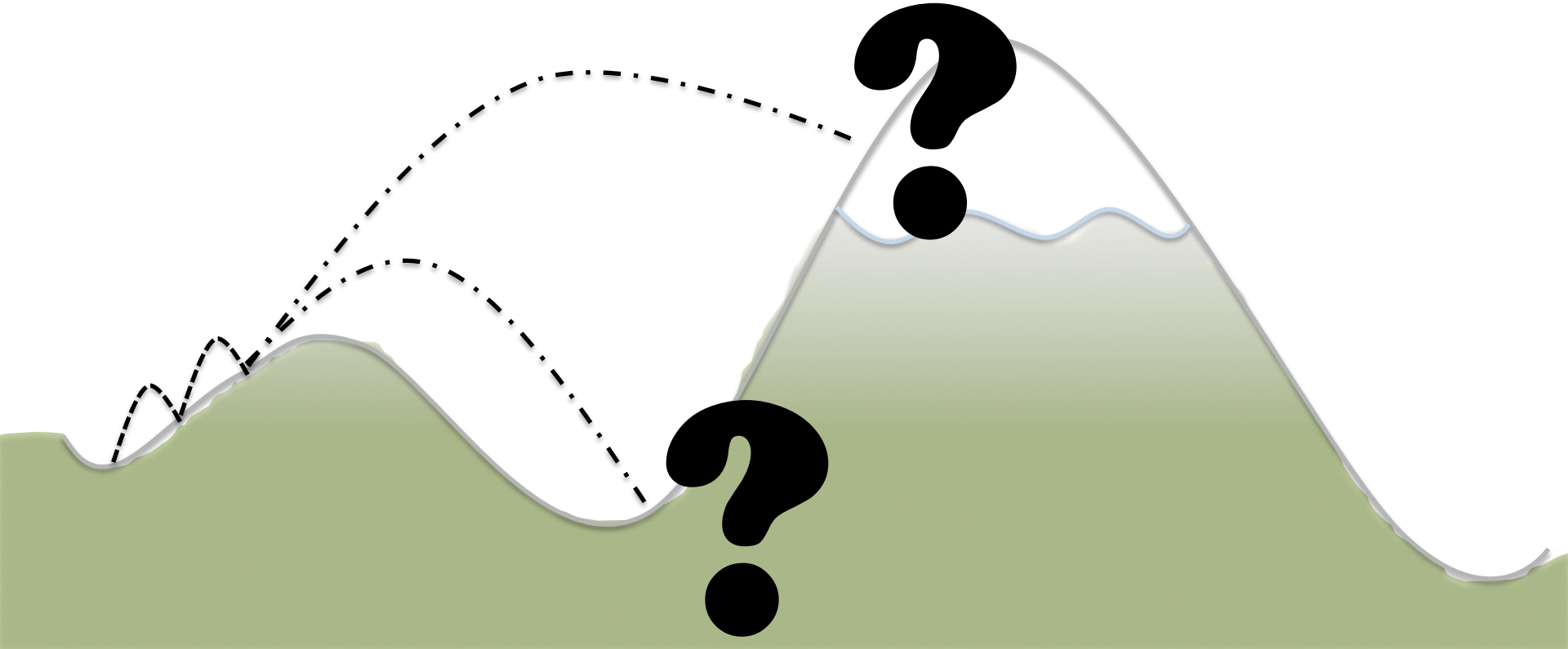
Making a big change all at once



Making a big change incrementally



Making the big leap



we always think it’s a huge leap **forward**

(but we’re not always right)

</how we got here>

<what to do right now>

Introduction to Pinterest

Find Friends

Invite Friends

I've often wondered how this could happen. How could a place so carefully, masterfully crafted be abandoned. Forgotten.

7 6

Found in
Abandoned Mansions

Little Italy, New York...I've walked the streets of Little Italy:ceeanne.

by camgraphe

1

Found in
Italy

Better late than never.

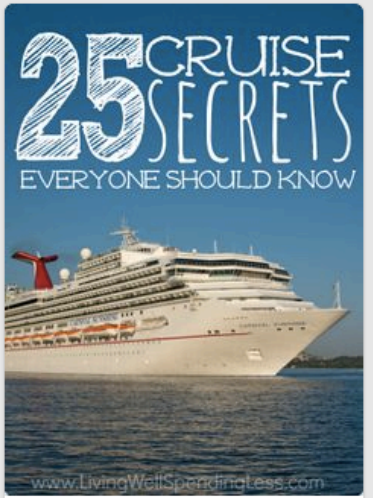
5

Found in
Typography

Modern Architecture Rules! Chris Pardo Designs

3

Found in
Modern Architecture



from Living Well Spending Less®

25 Cruise Secrets Everyone Should Know

Want to get away? Cruising is not only relaxing, it is also one of the most cost-effective ways to travel! These 25 insider secrets can help you find the best deals, discover little-known tips & tricks, and help you make the most of your next cruise vacation.

10 2

Found in
Travel

Website redesign

- Rewrite entire website
- Modular design for faster development
- Consume our API

- Complete visual redesign

Jellypop Pinkie

Anna Jackson-Jones onto American Eagle

Map ornaments from vacations, put the dates on the back.

Betsy Waugaman via Laura Black Pate onto sewing projects

Christmas Cookies...

Sarah Ireland via Misi V. onto holidays

OLSENS ANONYMOUS MARY KATE MK OLSEN FASHION STYLE BLOG QUILTED LEATHER JACKET GREY SHIRT LONG BLOND HAIR WHITE CLUTCH BALENCIAGA GLADIATOR WEDGE SANDAL

Natalia Namdar via Krystle Morgan onto Fashion

Shadows

Jules Le Masson Fletcher via Brent Heeb onto Photography Insipiration

peanut butter blossom = yummy old favorite

Brandi Brent Bogart via Suzy Marchetti onto Christmas

Black

by iloniti

sarah mccoy via Gabrielle Meehan onto Awesome

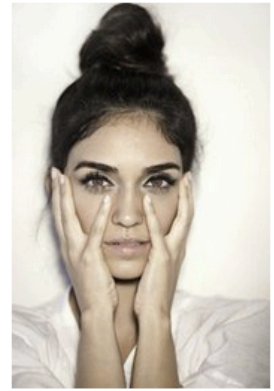
Vintage Cracker Jack prizes from the 60s and 70s

Sarah Bergman via Melanie Miller onto Plastic Trash!

Here's to having a great friend by your side as you loose your shit. Because going crazy alone is just not as much fun. @K . Doss

Danielle Rogers via Charlsa McCachren onto Funny things

Aislenn K onto Where to



Thick liner and lashes, bold brow and

from STYLECRAZE

Best Hairstyles For Shoulder Length Hair – My Top 10

Top 10 Hairstyles For Shoulder Length Hair

Alec Soto
Hair

Gorgeous Front Porch!

Lee Harrison-Acquaro
home

Cream Cheese Peppermint Dip | Safeway - Sweet and creamy peppermint dip served with chocolate-covered pretzels and chocolate graham crackers make an unexpected holiday dessert dip! Mix cream cheese, ricotta, powdered sugar and peppermint extract in a bowl on high for 2 minutes. Chill for 30 minutes before serving and then garnish with crushed peppermint candies or peppermint bark.

146 7

Safeway
Safeway Culinary Creations

Promoted Pin

"Feel the sunlight on my face You have brought me through this place." -Barlow Girl

Megan Bay
Clothing

from Trover

Goreme, Nevşehir, Turkey — by Ana Patrascu

Cappadocia, Turkey

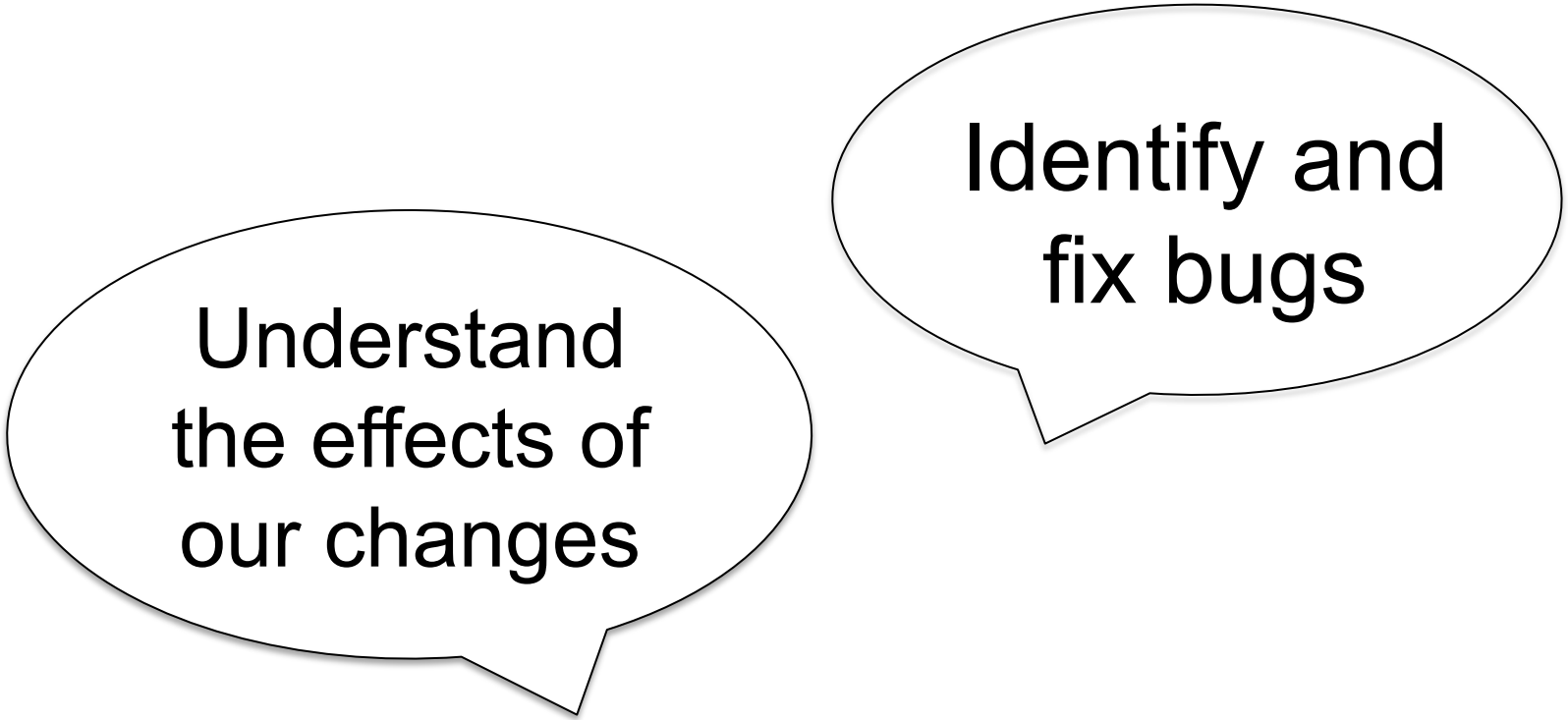
Holly Rosendahl
Istanbul

Why NOT run an experiment?

We know we'll ship – we can't go back to the old design now!

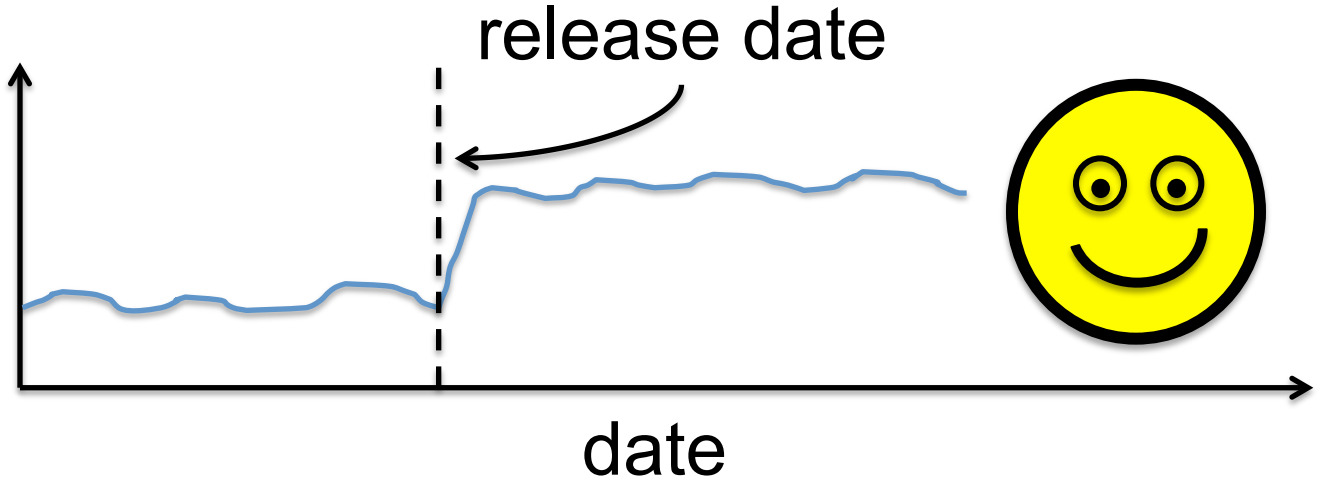
We've already told the press about it!

Why run an experiment?



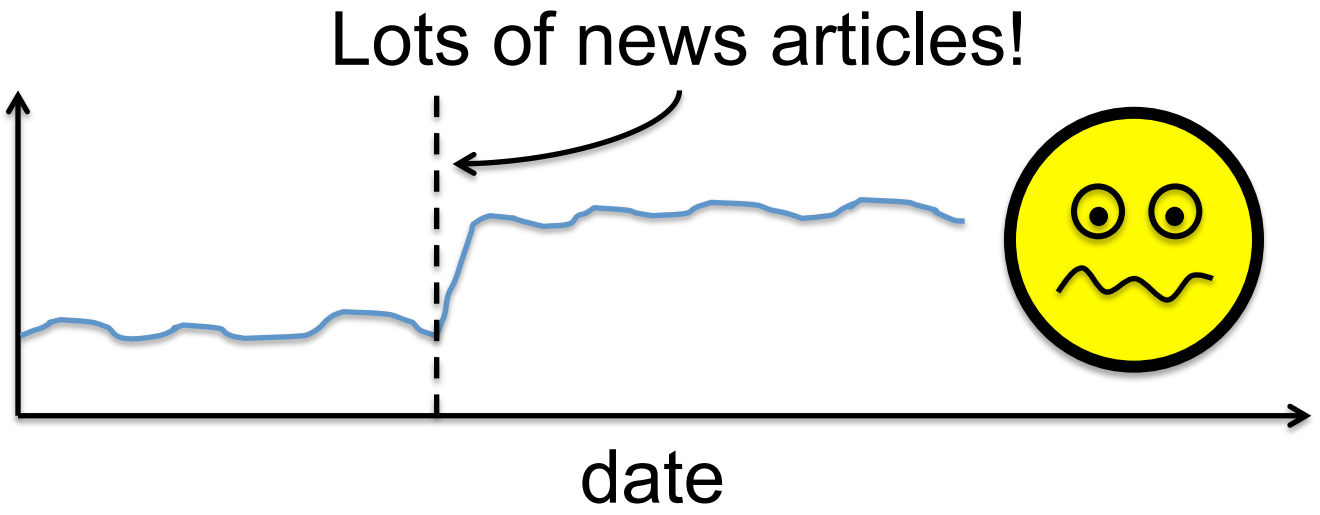
Option #1: Just ship it

- Release the new site.
- Watch metrics trend to evaluate success.



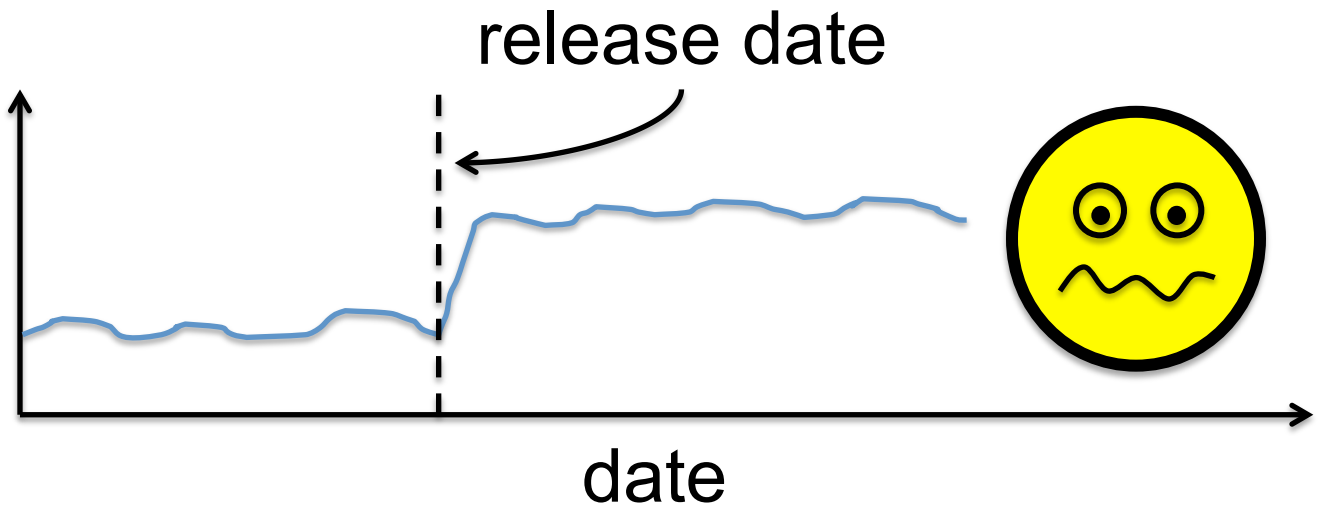
Option #1: Just ship it

- But: other factors may be responsible



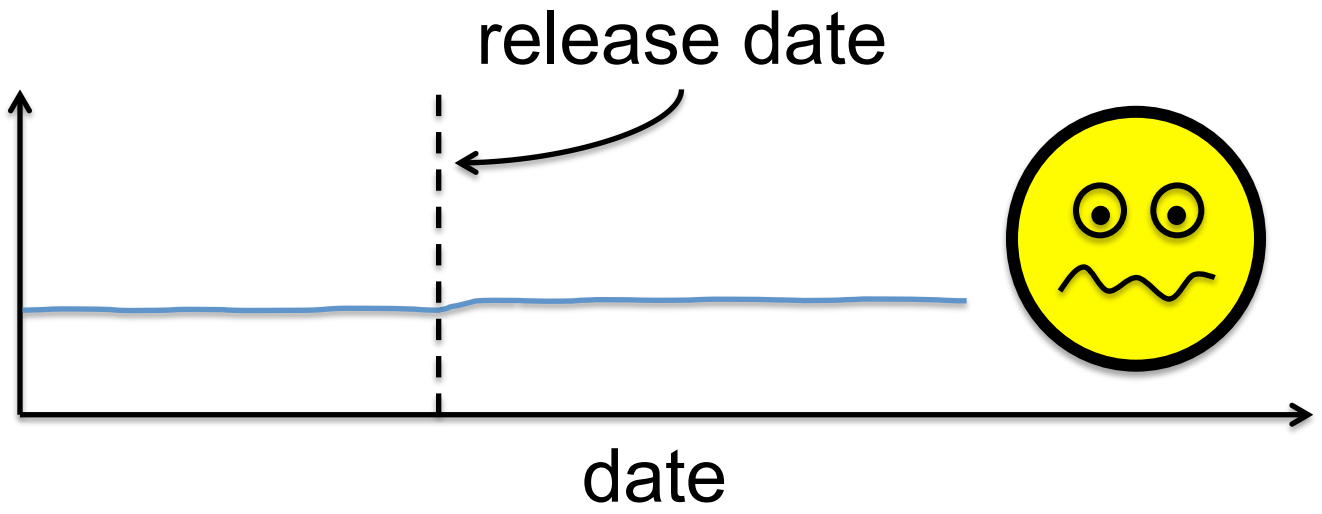
Option #1: Just ship it

- But: novelty effect could skew results



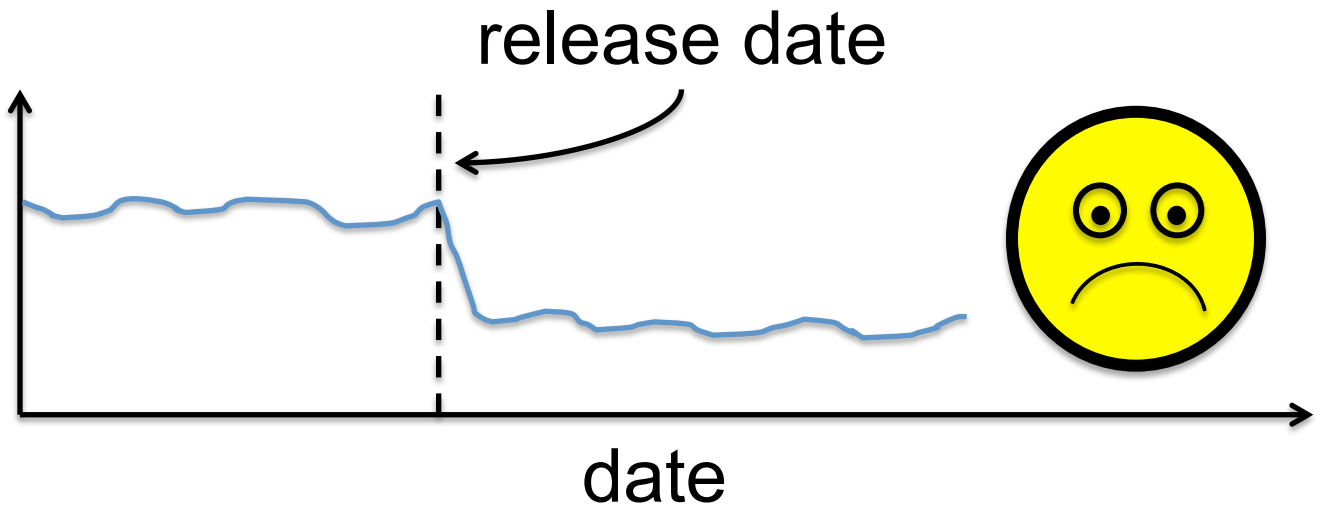
Option #1: Just ship it

- But: effect size may be too small



Option #1: Just ship it

- But: anything you find is too late!



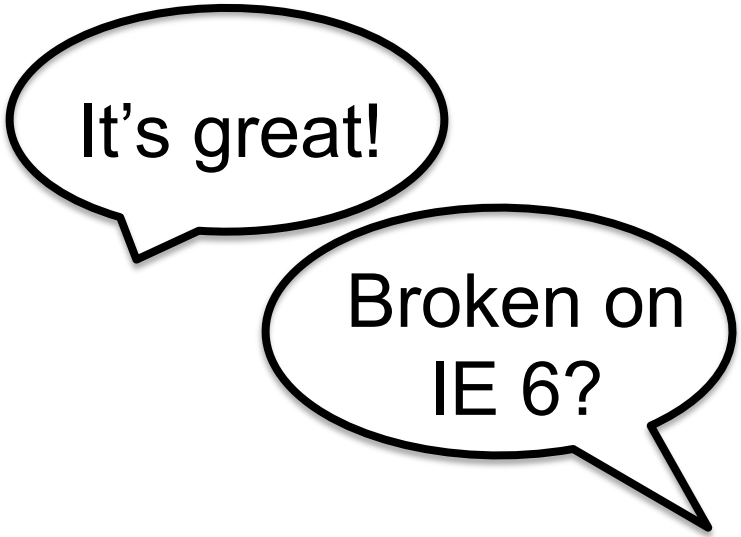
Option #1: Just ship it

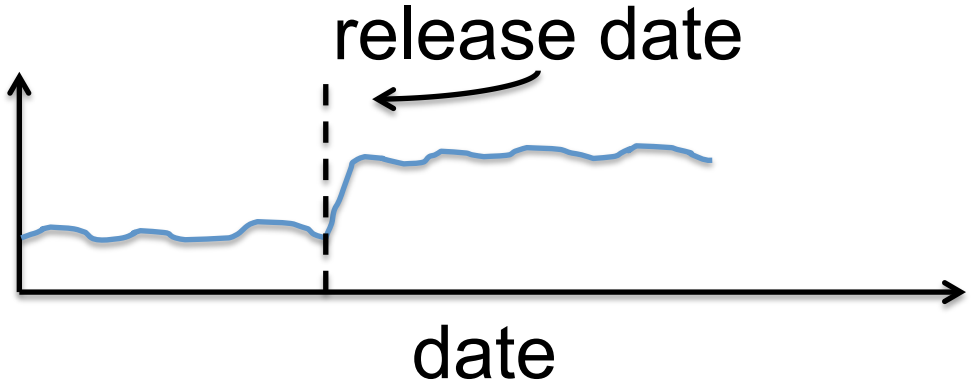
Options for the big release

- ~~Just ship it~~
- **Allow users to opt in**

Option #2: Let users opt in

- Let users opt in. Listen to their feedback.
- Watch metrics trends too.





Option #2: Let users opt in

Preview our new look

Pinterest

We're testing out a more visual design with bigger pins. We'd love to know what you think!
Learn more

No Thanks | Preview

Lucerne, Switzerland submitted by: itslikequack, thanks!

Abi Harris onto Places I Want To Visit

Glacier "Perito Moreno" in Argentina. (source:...

Abi Harris onto Places I Want To Visit

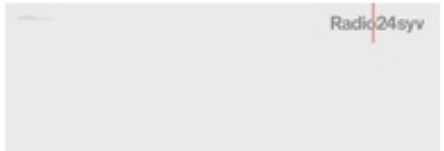
Brooklyn Bridge, NY
Travis Feldman Photo

Abi Harris on
To Visit

News See all

Cool Stuff

Your pin from **Random Wants & Desires** was repinned here:

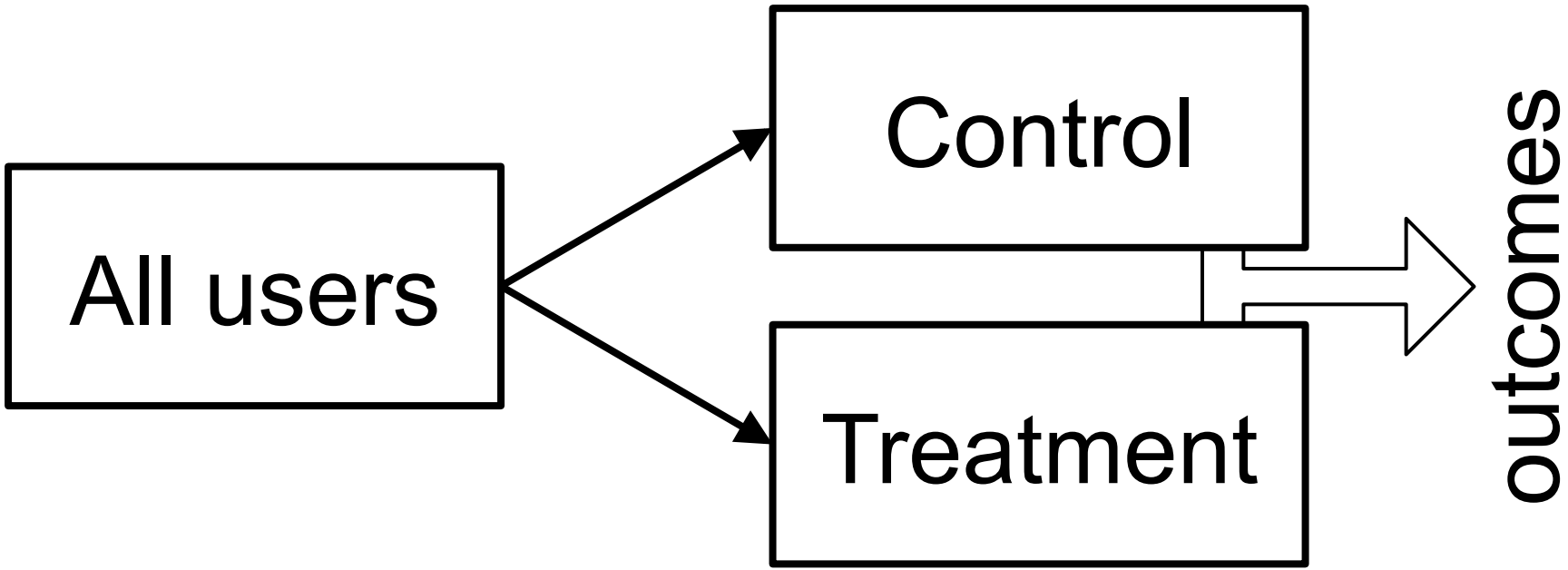


Option #2: Let users opt in

Options for the big release

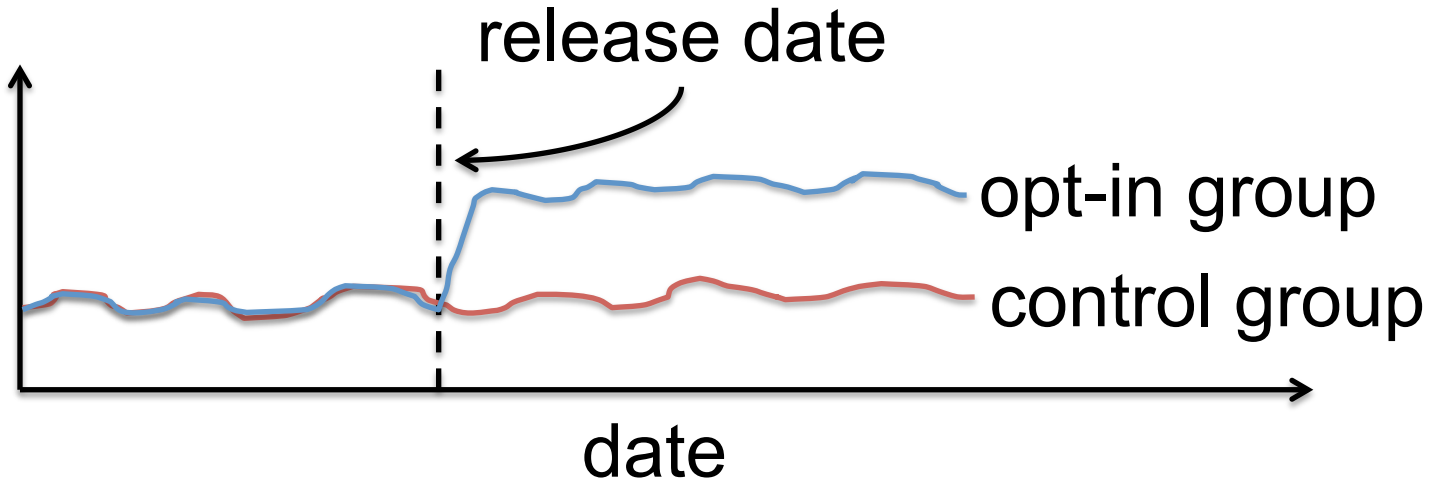
- ~~Just ship it~~
- (Allow users to opt in)
- **Run an opt-in experiment**

Option #3: Run an opt-in experiment



Option #3: Run an opt-in experiment

- Allow a random group of users to opt in
- Compare metrics to a control group



Challenges of opt-in: dilution

- Dilution decreases signal : noise ratio

a person in your experiment

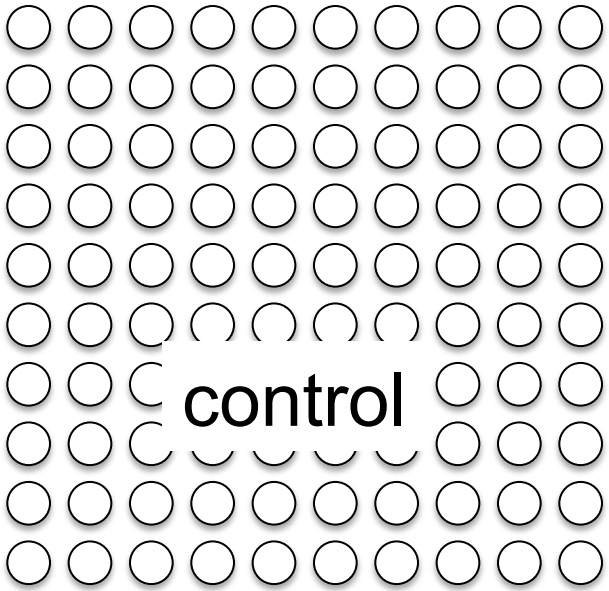
Challenges of opt-in: dilution

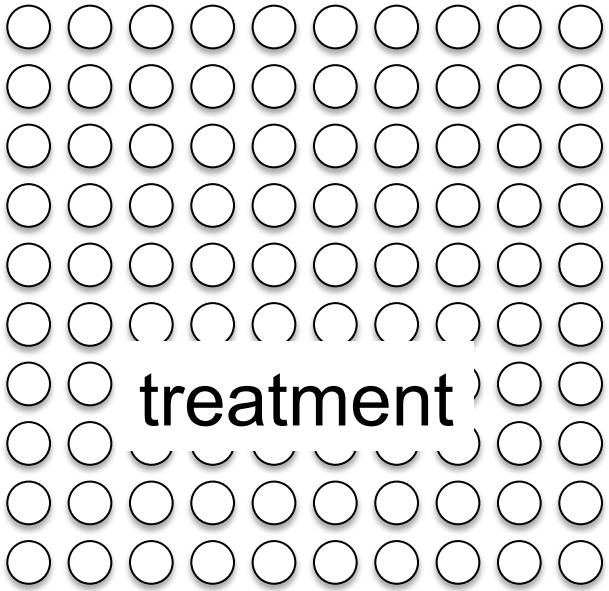
- Dilution decreases signal : noise ratio

a happy person in
your experiment

Challenges of opt-in: dilution

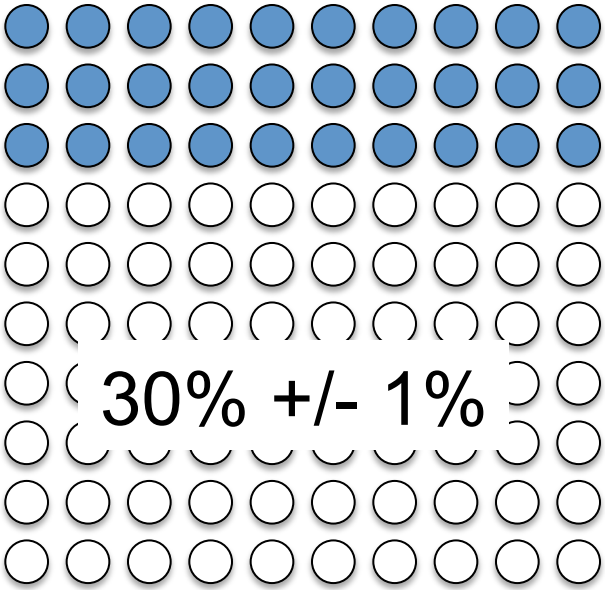
- Dilution decreases signal : noise ratio

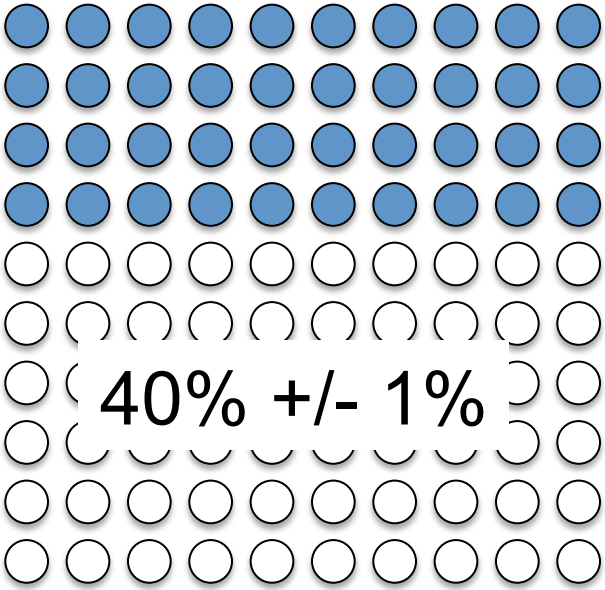




Challenges of opt-in: dilution

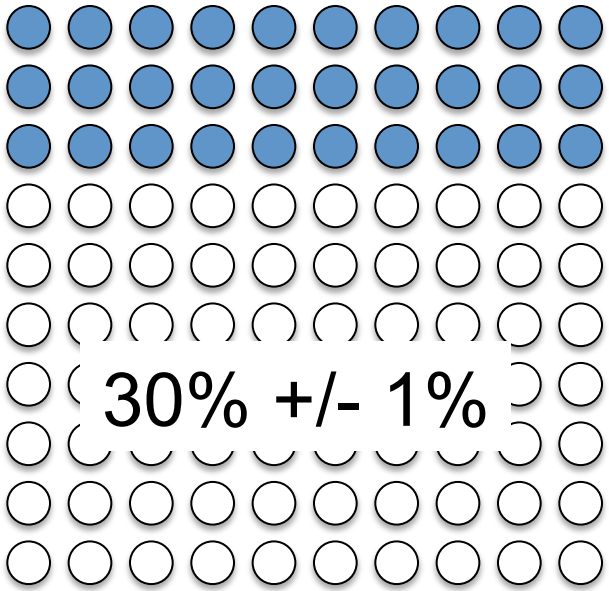
- Dilution decreases signal : noise ratio

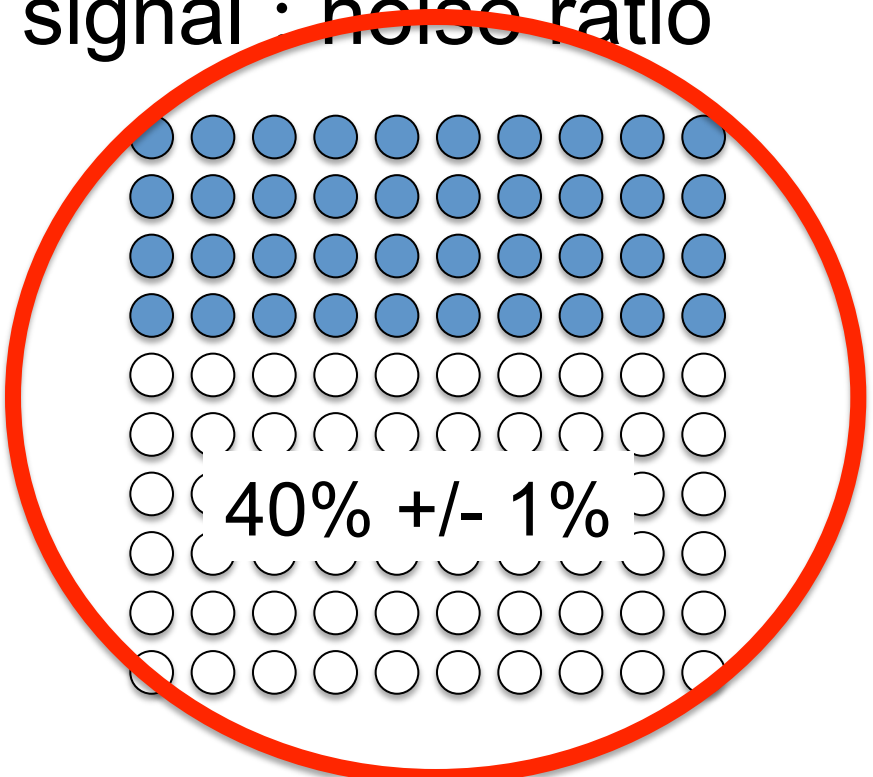




Challenges of opt-in: dilution

- Dilution decreases signal : noise ratio





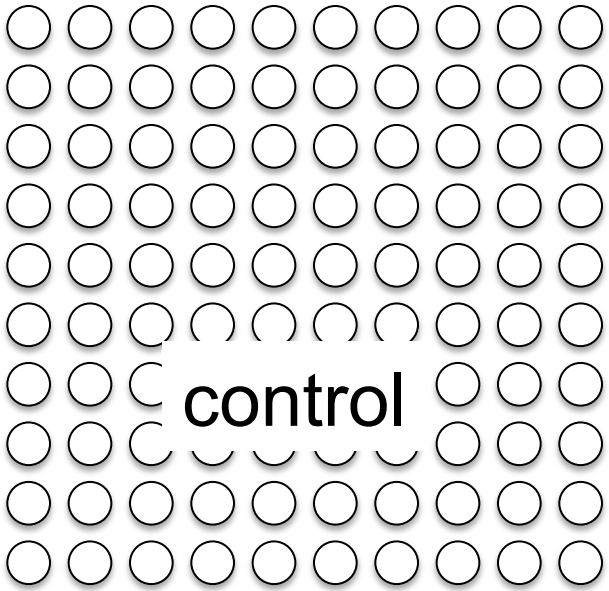
Challenges of opt-in: dilution

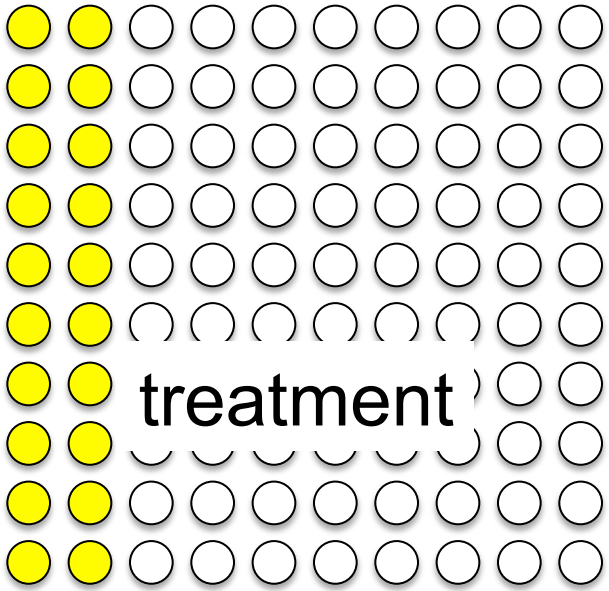
- Dilution decreases signal : noise ratio

a person who
opted into the new
website

Challenges of opt-in: dilution

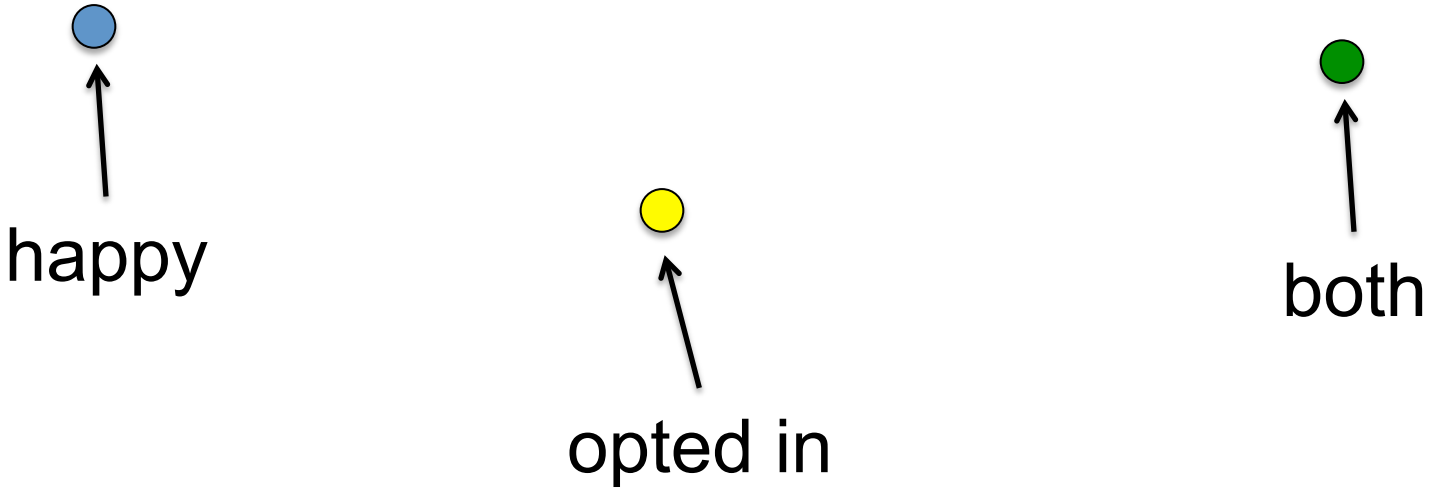
- Dilution decreases signal : noise ratio





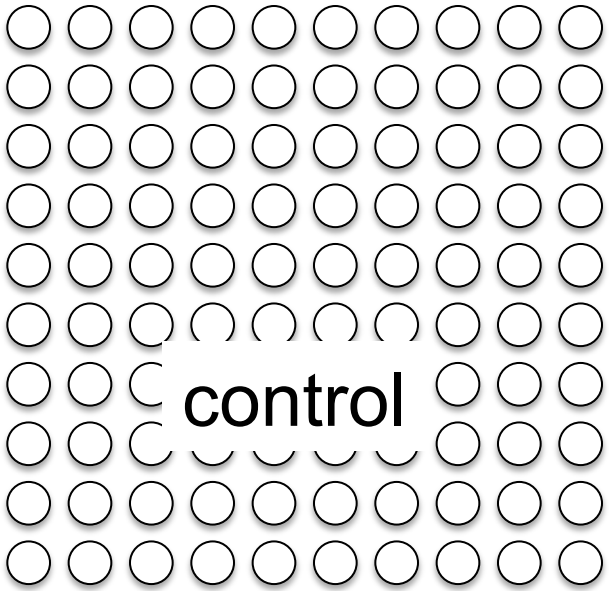
Challenges of opt-in: dilution

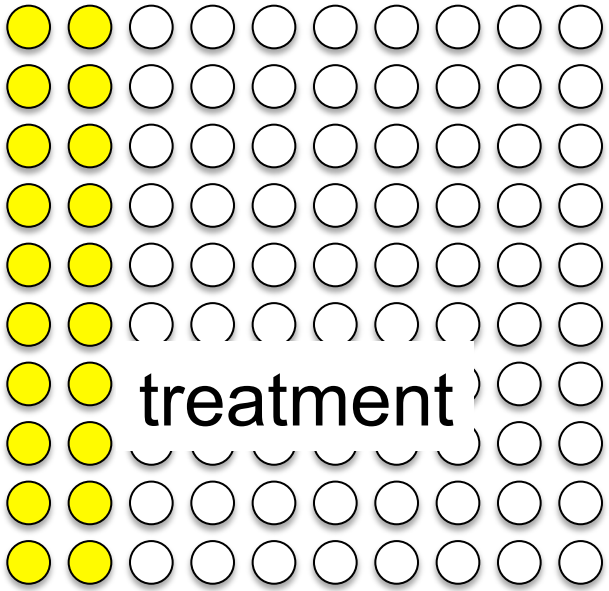
- Dilution decreases signal : noise ratio



Challenges of opt-in: dilution

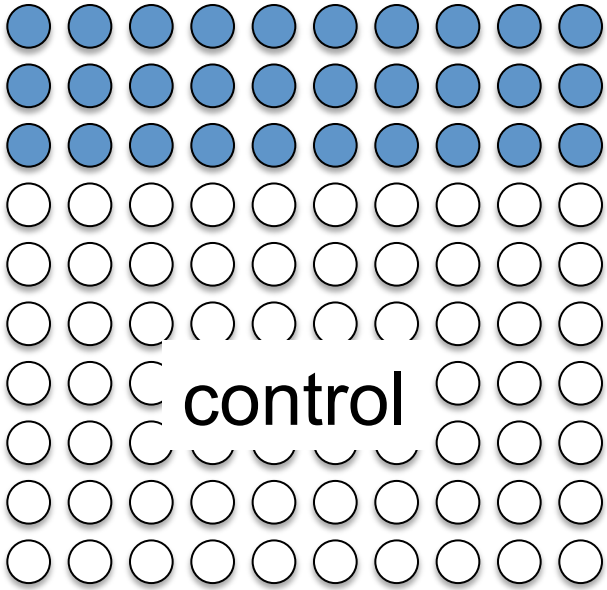
- Dilution decreases signal : noise ratio

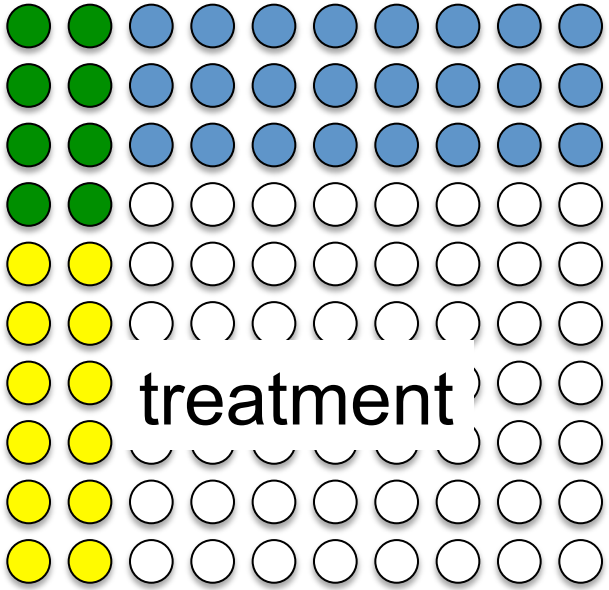




Challenges of opt-in: dilution

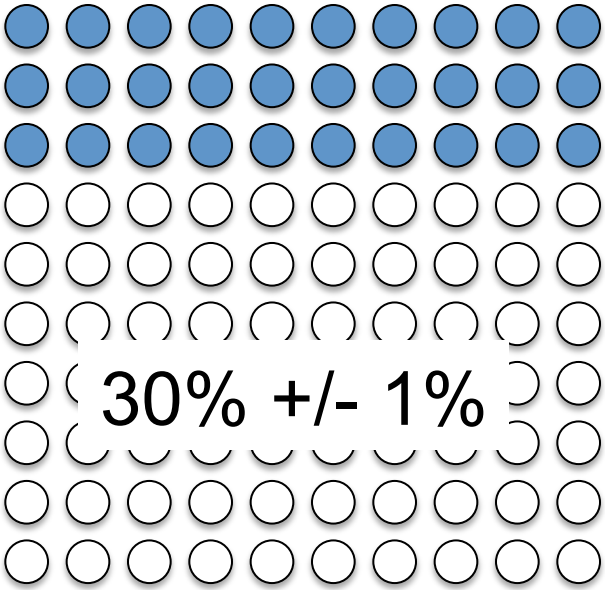
- Dilution decreases signal : noise ratio

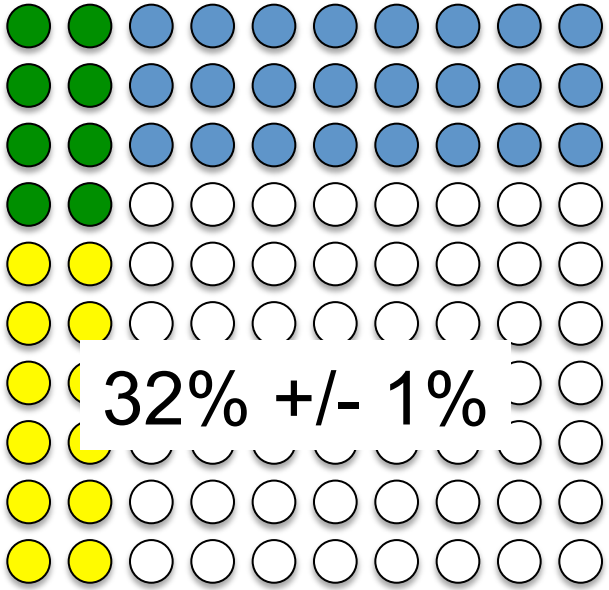




Challenges of opt-in: dilution

- Dilution decreases signal : noise ratio

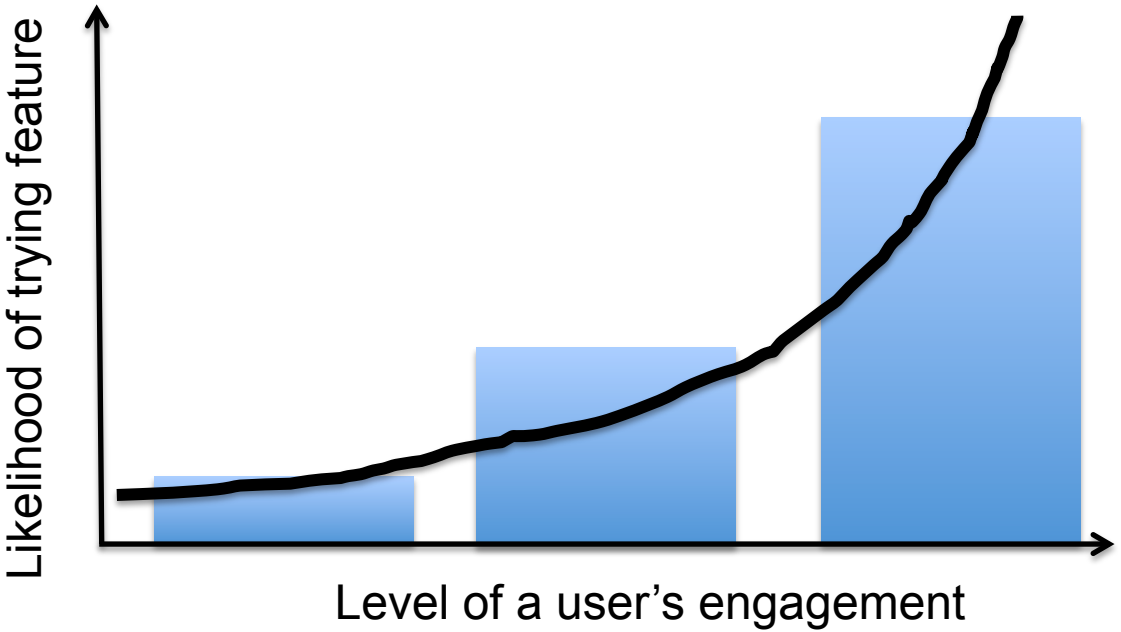




dilution of results =
may not be able to measure

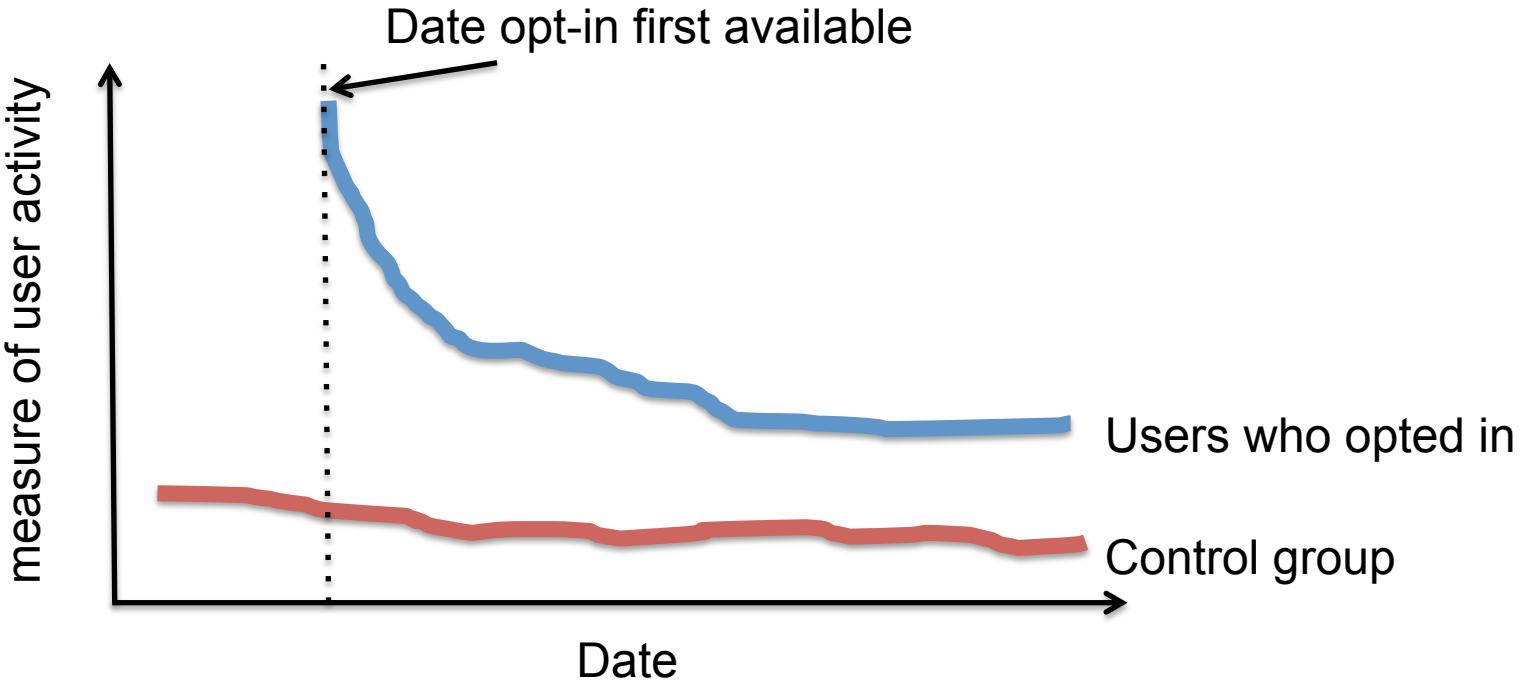
Challenges of opt-in: selection bias

- Engaged = more likely to try new features



Challenges of opt-in: selection bias

- Users who opt in differ from those who don't



Challenges of opt-in: selection bias

- Users who opt in differ from those who don't

Option #3: Run an opt-in experiment

Data is skewed by:

- Dilution
- Selection bias

Options for the big release

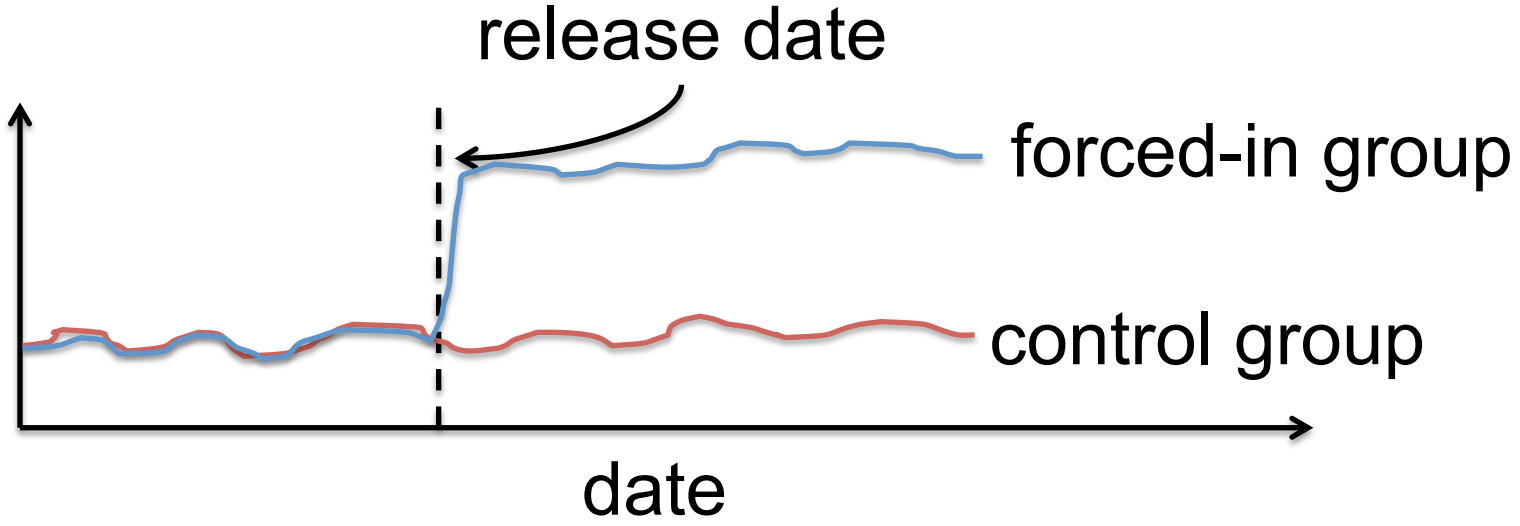
LEARN MORE

- ~~Just ship it~~
- (Allow users to opt in)
- ~~Run an opt-in experiment~~
- **Run a forced-in experiment**

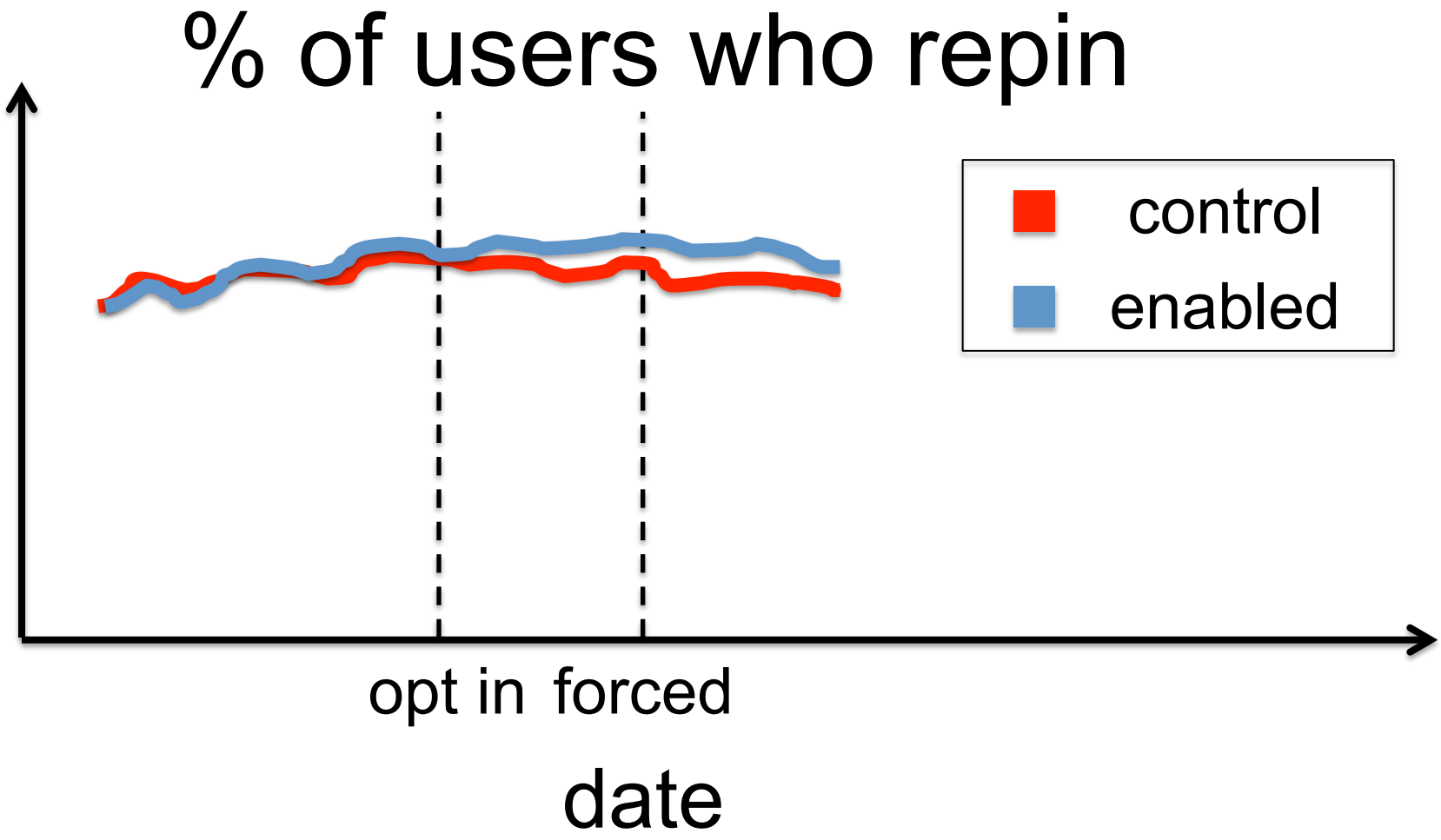
GO FASTER

Option #4: Run an forced-in experiment

- Force some users into the new design
- Compare metrics to a control group



April 3, 2013


% of users who repin
control
enabled
opt in
forced
date

“It’s not an experiment if you know it’s going to work.”

– Jeff Bezos

“My experiment will be hugely successful.”

– every feature owner ever

even optimists need a contingency plan

should we ship today?

how about now?

or now?

DENZEL LAUNCH

Pin creates, comments and user / board follows are still down. Everything else is up or neutral.

xx%

% web users in Denzel

EXISTING USERS

Pin create	↓
Board follow	↓
User follow	↓
Comments	↓
Likes	↑
Click through	↑
Board create	↑
Repin	↑

NEW USERS

Board follow	↓
User follow	↓
Pin create	↓
Board create	↑
Got repinned	↑
Repins	↑
Likes	↑
Return visits	↔

April 10, 2013

DENZEL LAUNCH

Pin creates, comments and user / board follows are still down. Everything else is up or neutral.

xx%

% web users in Denzel

EXISTING USERS

Pin create

User follow

Comments

Board create

Repin

NEW USERS

Likes

Return visits

bad enough to block shipping?

DENZEL LAUNCH

Pin creates, comments and user / board follows are still down. Everything else is up or neutral.

xx%

% web users in Denzel

EXISTING USERS

- Pin create
- Board follow
- User follow
- Comments
- Likes
- Click through
- Board create
- Repin

NEW USERS

- Board follow
- User follow
- Pin create
- Board create
- Got repinned
- Repins
- Likes
- Return visits

how bad is too bad?

DENZEL LAUNCH

Pin creates, comments and user / board follows are still down. Everything else is up or neutral.

xx%

% web users in Denzel

EXISTING USERS

Pin create

Board follow

User follow

Comments

Likes

Click through

Board create

Repin

NEW USERS

Board follow

User follow

Pin create

Board create

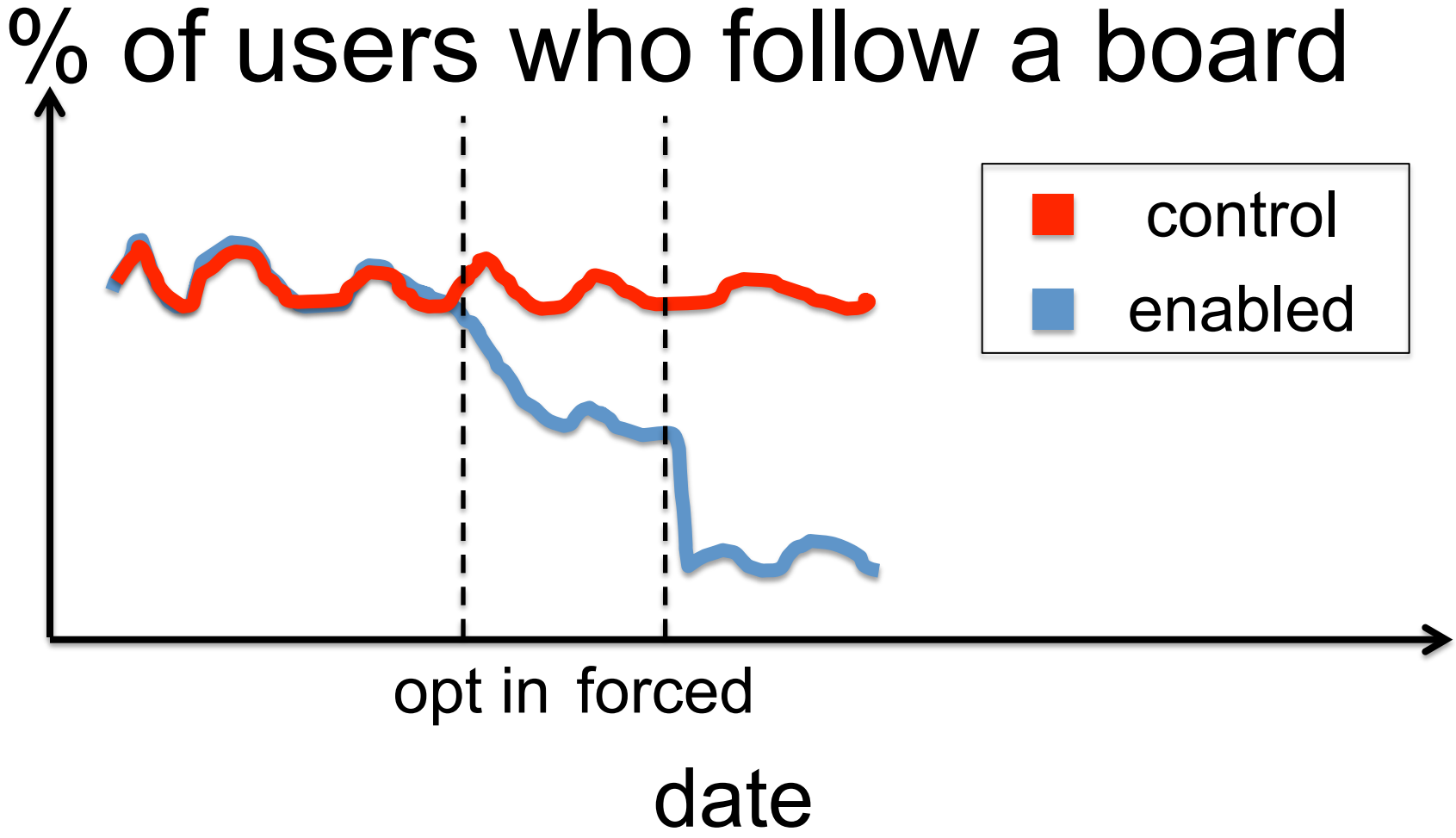
Got repinned

Repin

Likes

Return visits

we need to fix this before we ship!


% of users who follow a board
control
enabled
opt in
forced
date

[redacted]@pinterest.com> Unsubscribe 4/5/13

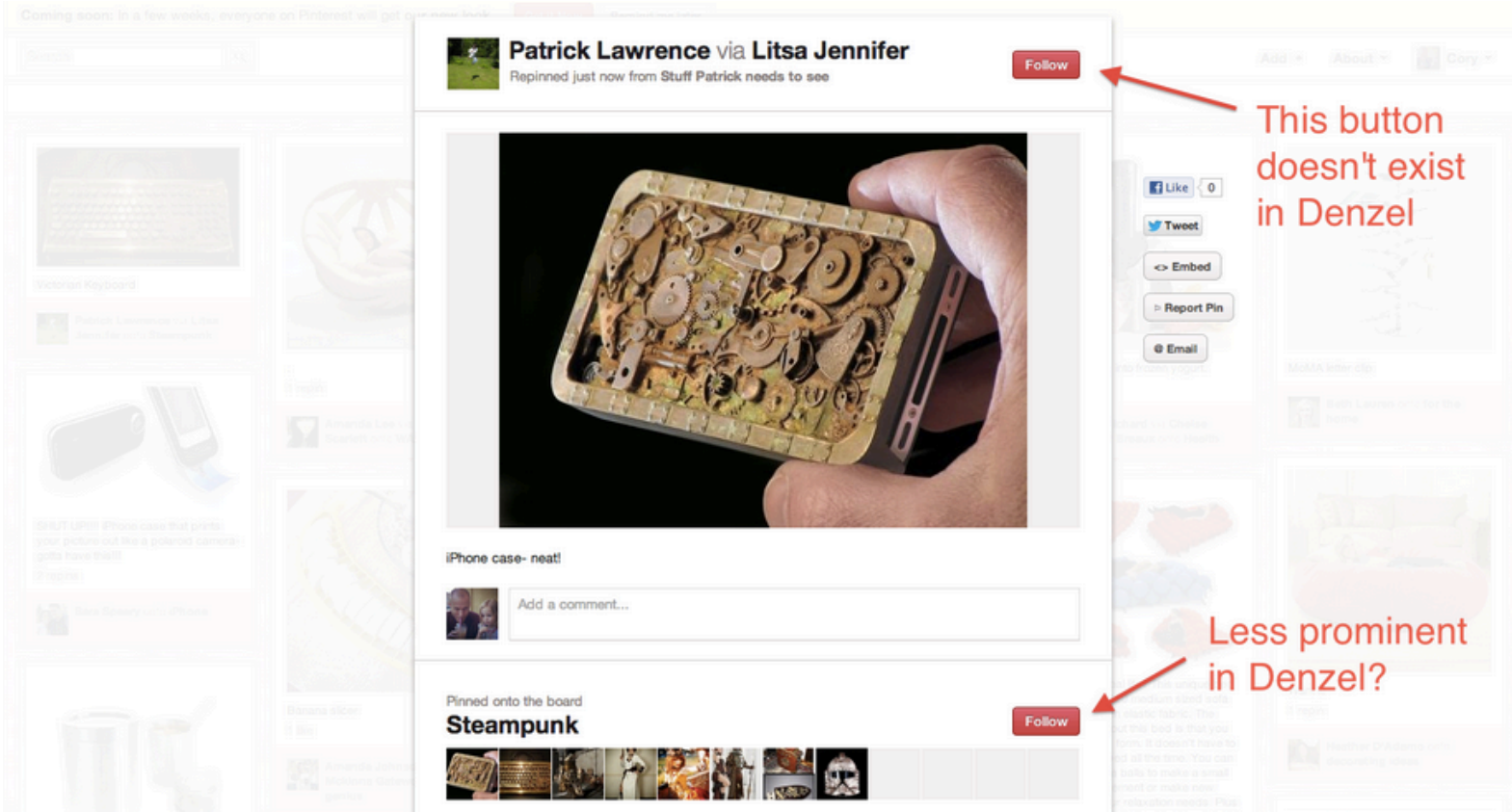
to team

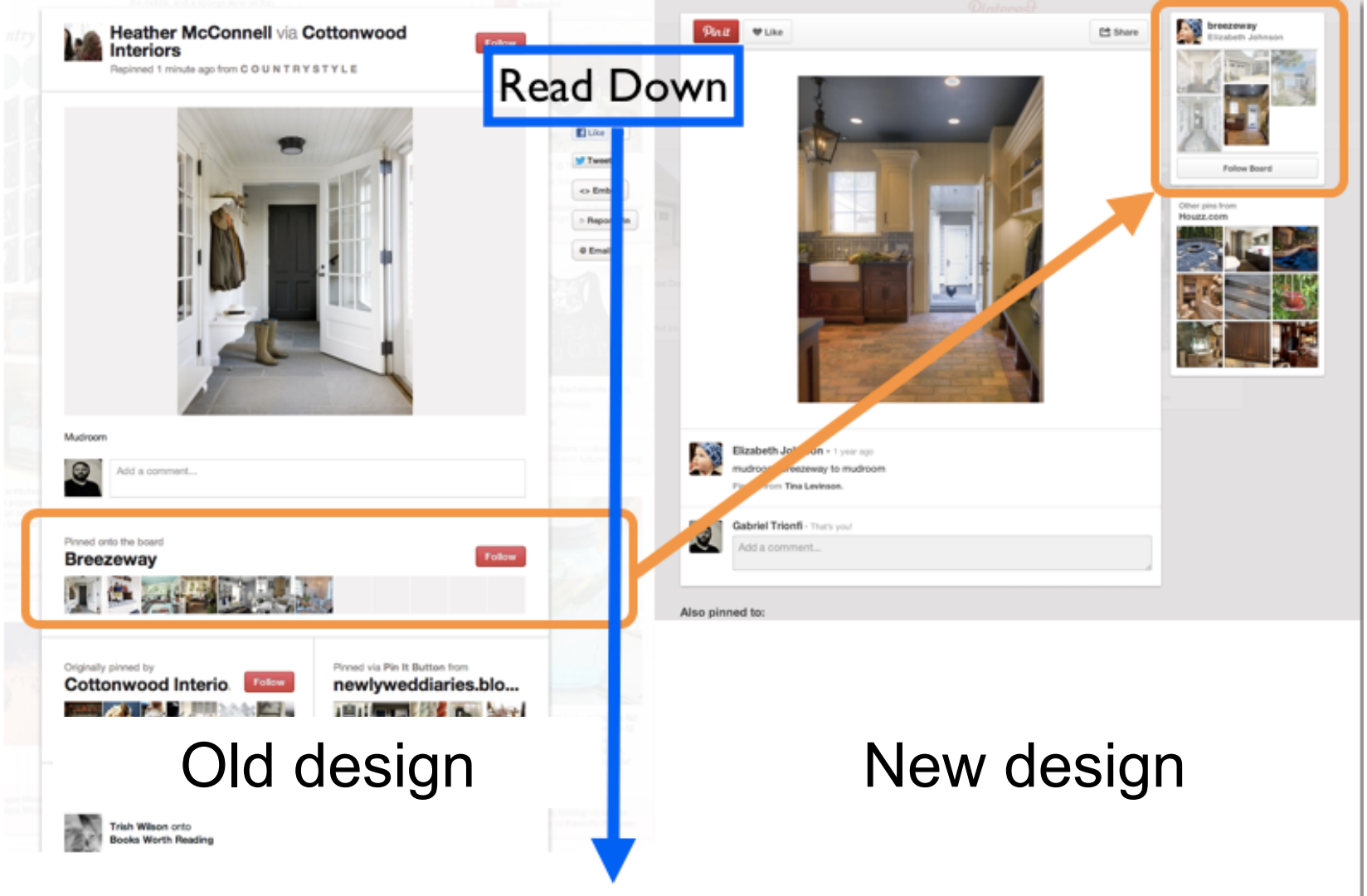
As Ben mentioned in Q&A, initial data show that most of Denzel's key metrics are either neutral or positive except for board follows. Because the most reliable data we have is from the "new user" experiment and the "forced in existing user" experiment (both of which began this week), our data is still very early and we should be prepared for some fluctuation.

That aside, there are a few theories that we have as to why board follows are down. For example:

- **Closeup.** The board follow button on closeup in Classic is red and under the pin image whereas in Denzel the board follow button is gray and off to the right of the pin image (bottom part of the board preview module).
- **Repin dialog.** The repin dialog in Denzel doesn't have a board recommendation on repin success like Classic does.
- **Harder to find users to follow.** Finding people to follow obviously increases user follows, but it also increases board follows.
    - The "**pinned via**" meta data on the pin in grid helped people find pinners to follow but it's not in the Denzel pin grids.
    - Traffic to **category feeds** may be down due to the new navigation. People find new pinners to follow from the category feeds.

Theories about board follows

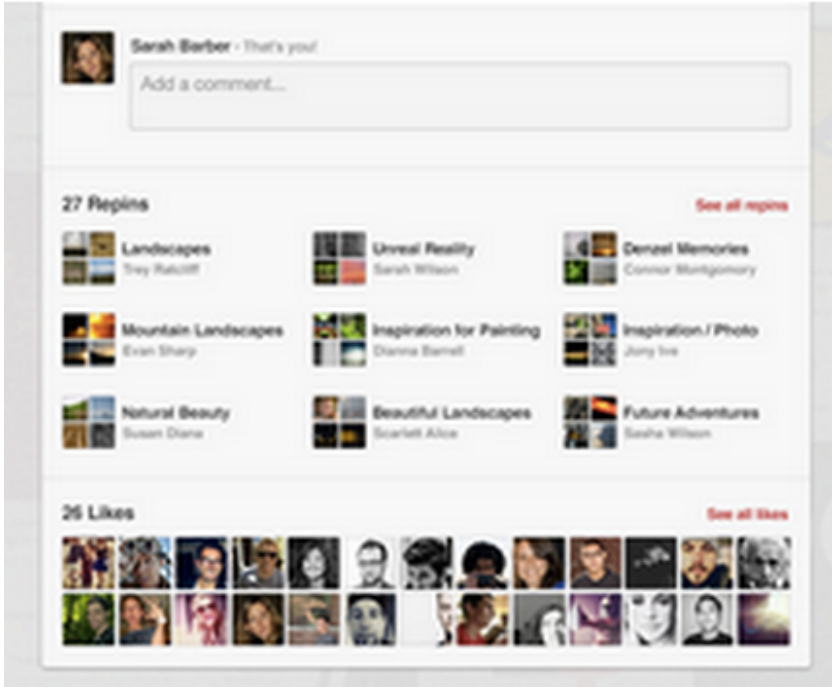



Heather McConnell via Cottonwood Interiors
Repinned 1 minute ago from COUNTRYSTYLE
Mudroom
Add a comment...
Pinned onto the board
Breezeway
Follow
Originally pinned by
Cottonwood Interio
Pinned via Pin It Button from
newlyweddiaries.blo...
Trish Wilson onto
Books Worth Reading
Read Down
Pin it
Like
Share
Follow Board
Other pins from
Houzz.com
Elizabeth Johnson
Gabriel Trionfi
Add a comment...
Also pinned to:
Old design
New design

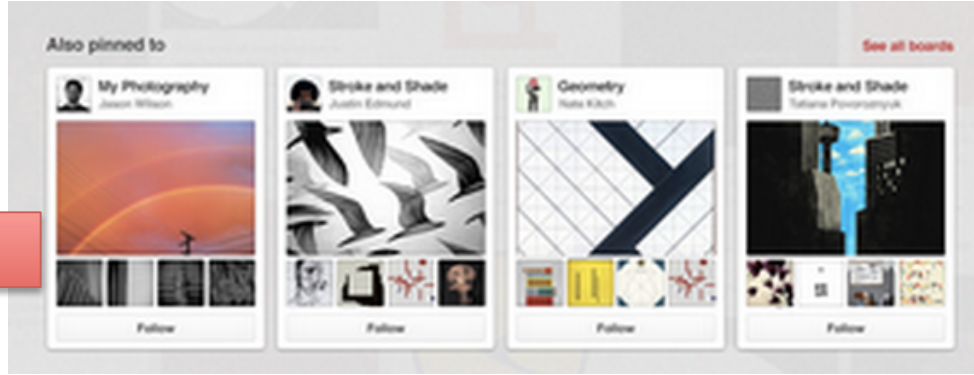
“It doesn’t matter how beautiful your theory is … If it doesn’t agree with experiment, it’s wrong.”

– Richard Feynman

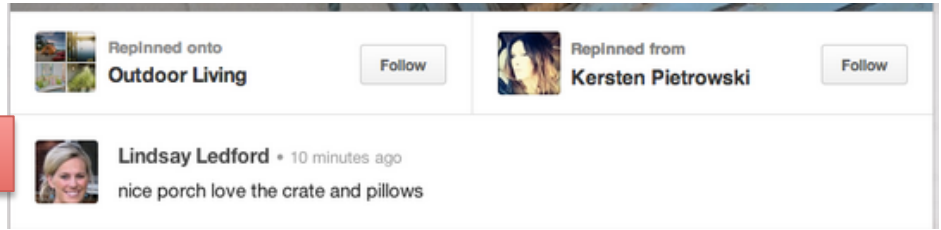
Show likes and repins



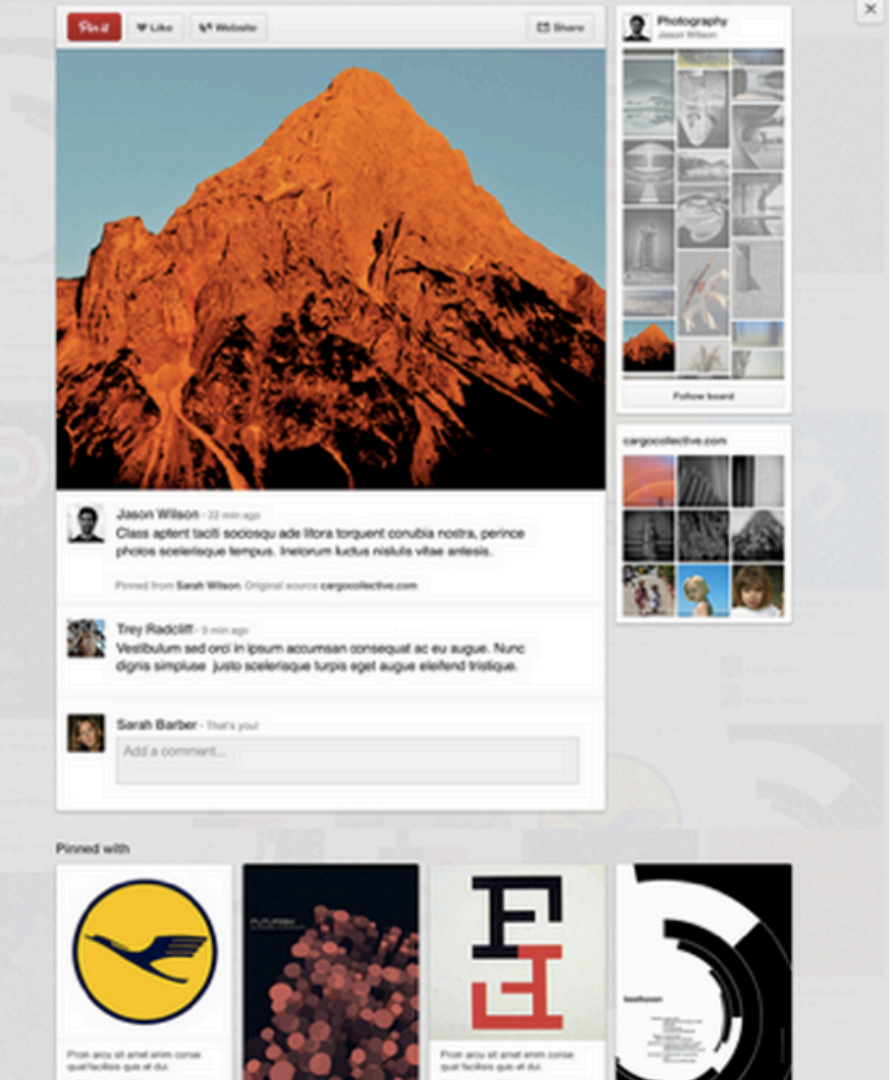
Show related boards



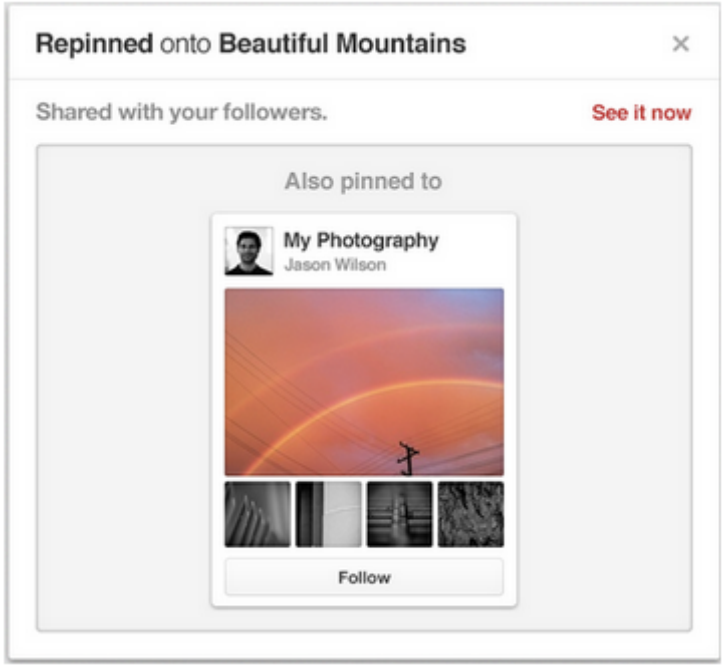
Show board & user under closeup

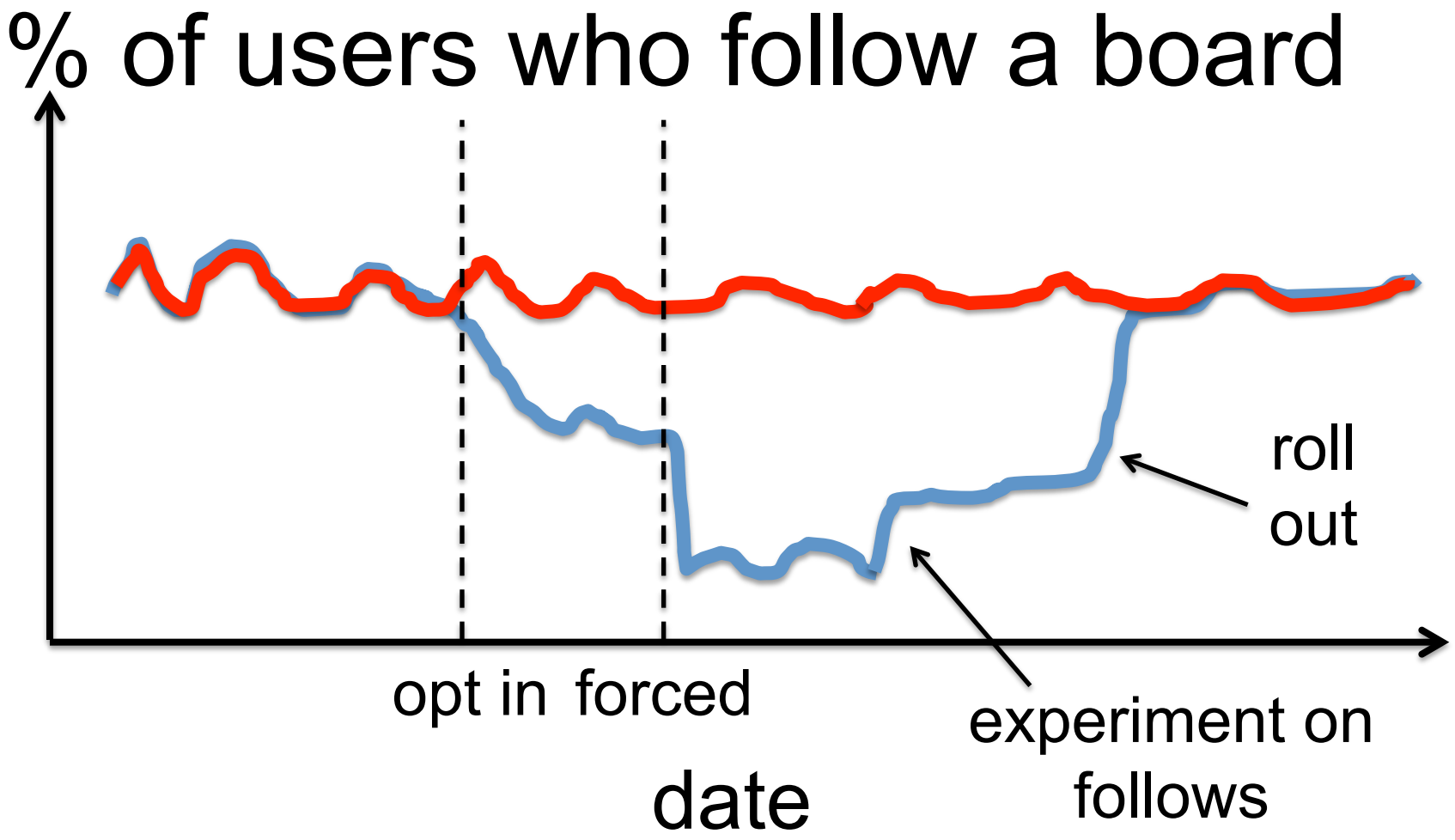


Red follow button on sidebar

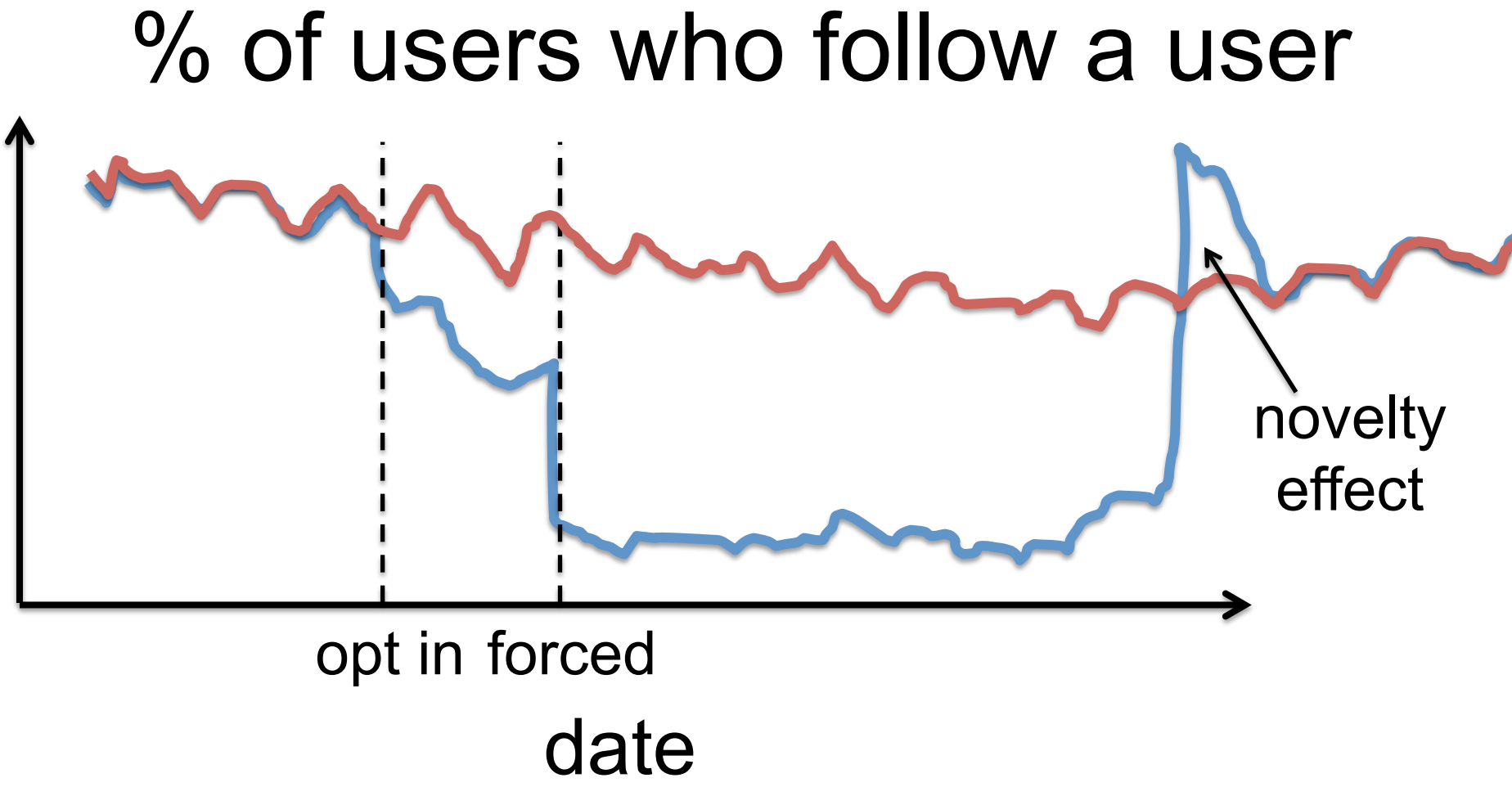


Board suggestion after repin




% of users who follow a board
roll
out
opt in
forced
experiment on
follows
date

should we ship now?


% of users who follow a user
novelty
effect
opt in
forced
date

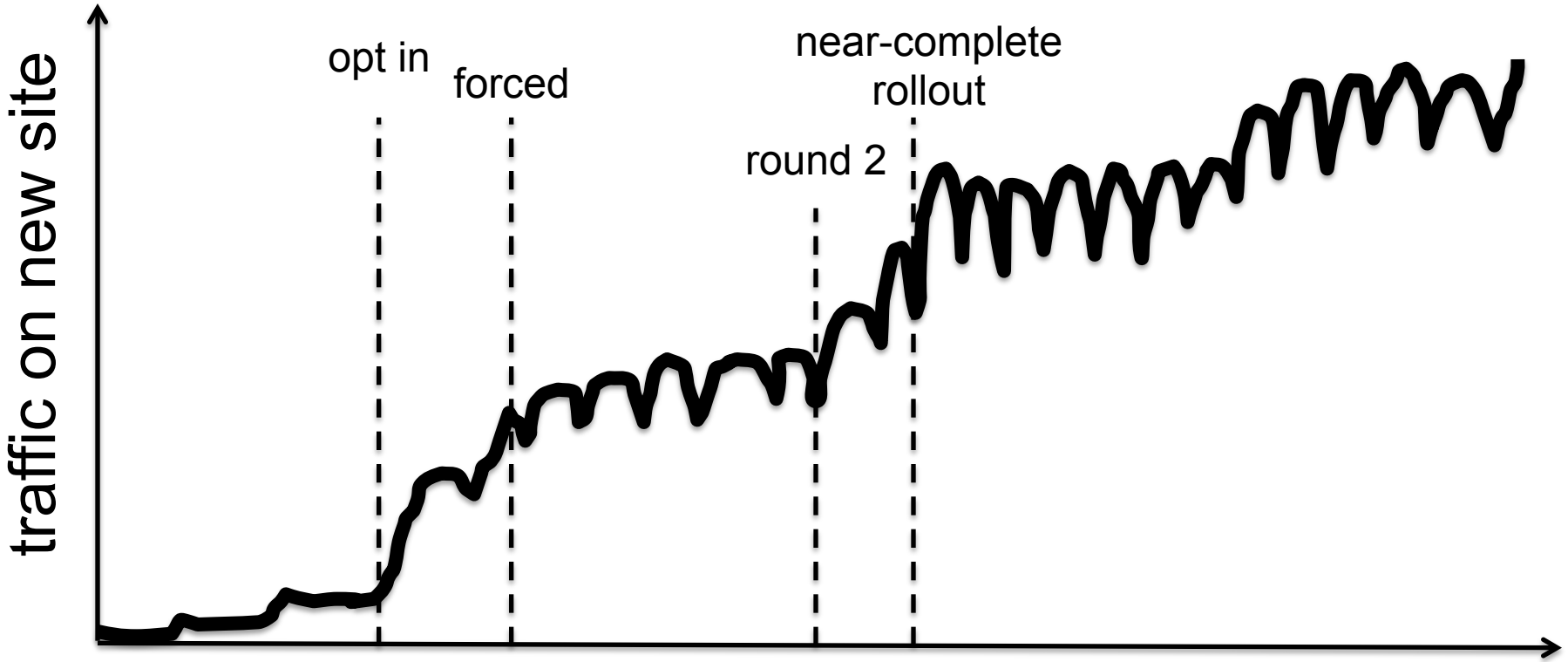
should we ship now?

a controlled experiment
gives you
disaster avoidance

a controlled experiment
gives you
accurate forecasting

a controlled experiment
gives you
confidence

Roll out with confidence



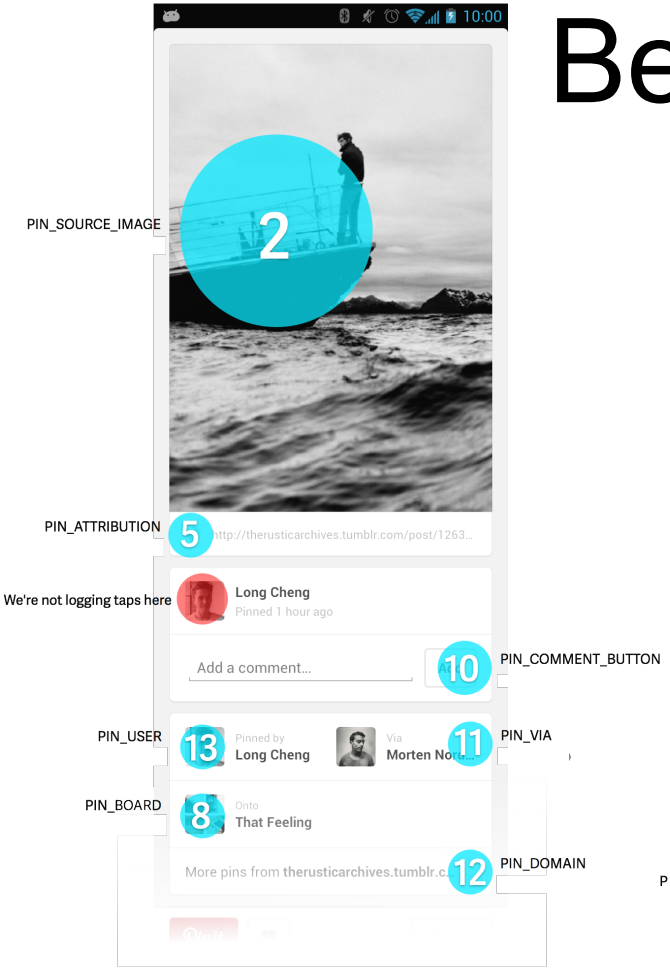
AND THEY LIVED

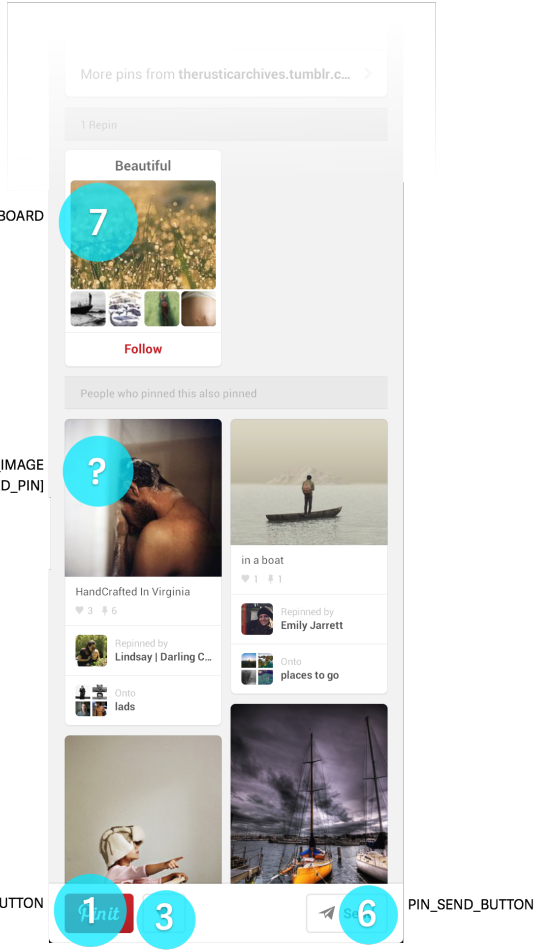
Happily ever after

Epilogue

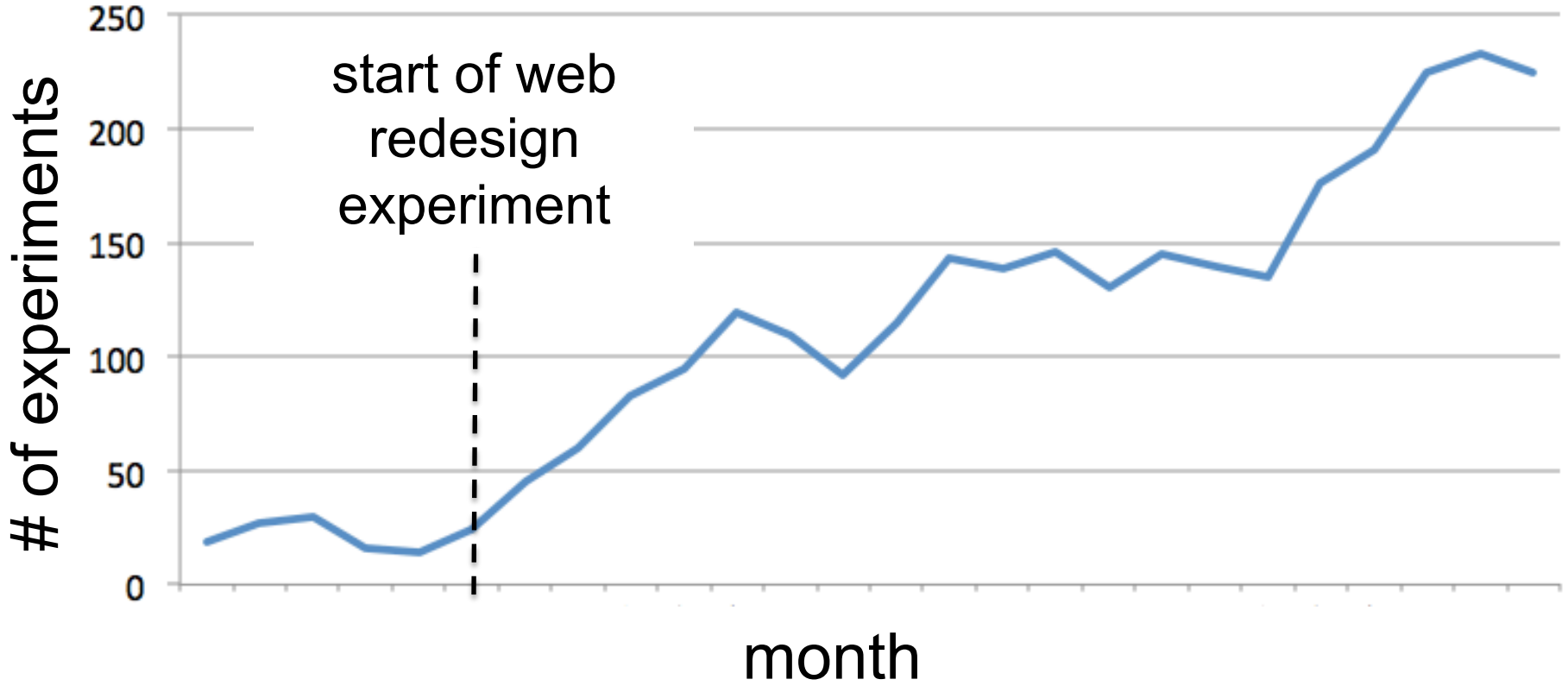
Better logging

where
users
click

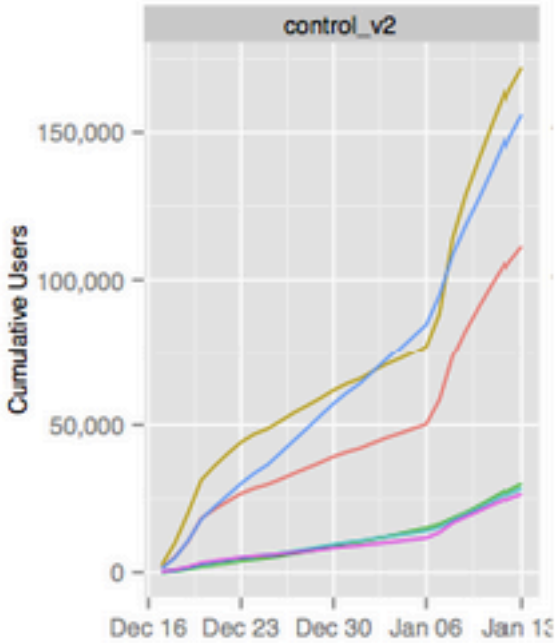


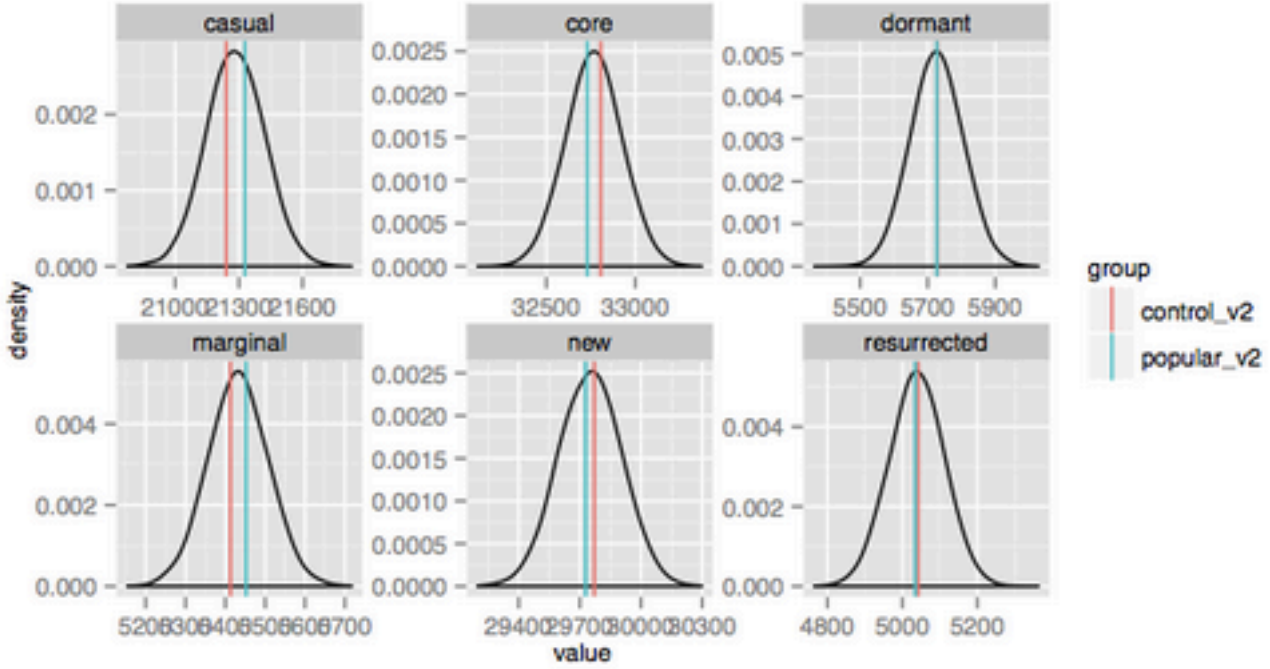


More experimentation

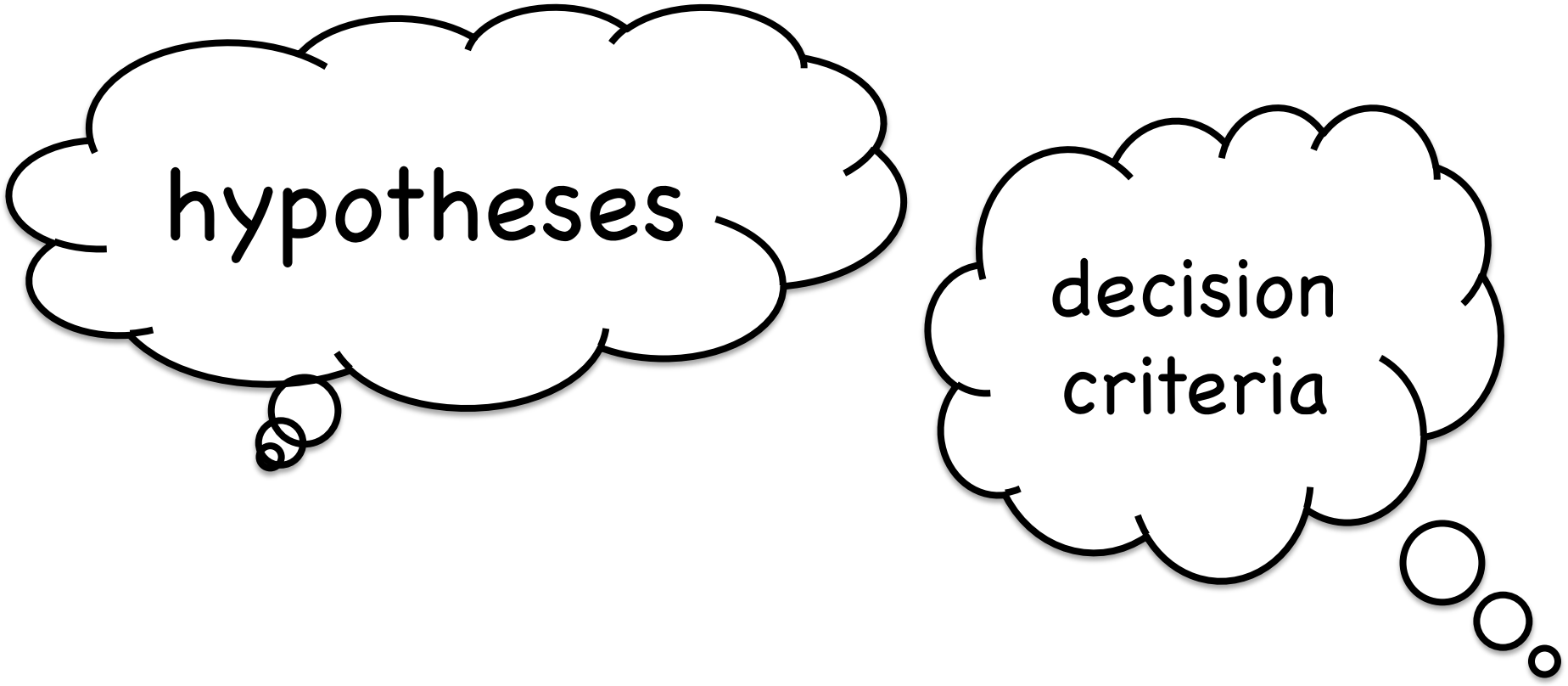


Automated diagnostic tools





Experiment process



Lessons from our grand experiment

be prepared:
your org will propose big changes

help your company decide when to go big

consider the tradeoffs:
less time testing, but more confusion

consider the tradeoffs:
incremental change = fast learning

help your company learn from big changes

measure the big bets:
find mountains, not valleys

measure the big bets:
run a controlled trial

measure the big bets:
avoid the opt-in trap

help your company
become more data-driven

learn from your big bets:
value of experiments & data

help your company think

Lessons from our grand experiment

be prepared: your org will propose and ship big things that are hard to learn from

understand the tradeoffs of rejecting incremental change

run an experiment to learn, and to avoid the valleys

randomize all users and measure predefined success criteria

expect high cost and high stakes, but extract high benefit

Thanks!

Questions?

References

- www.infoq.com/presentations/ab-testing-pinterest
- www.pinterest.com/burbsie11/ab-testing/
- about.pinterest.com/en/careers