MOZCON

2014

WELCOME!

ARCHITECTING GREAT EXPERIMENTS

Kyle Rush • Optimizely

@kylerush • kyle@optimizely.com

#MozCon

MEANS

WRITE IT DOWN

OR

TWEET IT

EXPERIMENT CONCEPTS

SIGNIFICANCE

The risk of encountering a false positive.

95% = out of 100 a/a tests, 5 will inaccurately report a difference.

POWER

The risk of encountering a false negative.

80% = out of 100 a/b tests, 20 winners will not be reported.

ONE-TAIL VS. TWO TAIL

One tail = is the variation better?

Two tail = is the variation better OR worse?

MDE

Minimum detectable effect

SAMPLE SIZE

How many subjects are in your experiment.

Always use a sample size calculator to calculate sample size before starting an a/b test.

bit.ly/VUBti8

SAMPLE SIZE CALCULATOR

bit.ly/SWR3YC

EXAMPLE

Absolute lowest MDE

CONVERSION RATE: 4%

MDE: 1%

POWER: 80%

SIGNIFICANCE: 5%

TAILS: ONE

2,972,435

Visitors per branch

EXAMPLE

Focus on time

1 month = 170,000 unique visitors

CONVERSION RATE: 4%

MDE: 6%

POWER: 80%

SIGNIFICANCE: 5%

TAILS: ONE

83,230

Visitors per branch

EXAMPLE

Small startup

1 month = 3,000 unique visitors

CONVERSION RATE: 4%

MDE: 45%

POWER: 80%

SIGNIFICANCE: 5%

TAILS: ONE

1,567

Visitors per branch

Sample size calculators tell you how many subjects,
but not which subjects should be in your experiment.

SAMPLING

- can be really hard
- week day vs. weekend traffic
- campaign vs. organic traffic
- returning vs. new visitors

WHERE SHOULD I TEST?

EXAMPLE

E-commerce website

Order before 1pm PST for **FREE Next Business Day** shipping on all **Clothing.** Learn More

24/7 Customer Service (800) 927-7671 | Help | Live Help | Log In or Register | My Account | My Favorites

Zappos.com POWERED by SERVICE®

Shoes, Clothing, Bags, etc. SEARCH

CLEARANCE SHOP NOW »

FREE SHIPPING & RETURNS

MY CART

SEARCH BY: *Size*, *Narrow Shoes*, *Wide Shoes*, *Popular Searches*

SHOES | CLOTHING | BAGS & HANDBAGS | AT HOME | BEAUTY | ACCESSORIES | SHOP BY... | WOMEN'S | MEN'S | KIDS' | ALL DEPARTMENTS

ALPHABETICAL BRAND INDEX # • A • B • C • D • E • F • G • H • I • J • K • L • M • N • O • P • Q • R • S • T • U • V • W • X • Y • Z

SHOP WOMEN'S
- Clothing
- Shoes
- Sandals
- Sneakers & Athletic Shoes
- Dresses

SHOP MEN'S
- Clothing
- Shoes
- Sneakers & Athletic Shoes
- Sandals
- Jeans

SHOP KIDS'
- Girls' Clothing
- Boys' Clothing
- Girls' Shoes
- Boys' Shoes

SHOP GIFT CARDS

SANDALS
SHOP NOW »

BACKPACKS
SHOP NOW »

ENTER FOR A CHANCE TO
WIN
A FREE PAIR OF SHOES

HOMEPAGE ASSUMPTIONS

1. Lots of traffic
2. Relatively few conversions

Let's estimate:

- 2.5% conversion rate
- 100,000 monthly unique visitors

RUN EXPERIMENT FOR 1 MONTH (100K VISITORS)

CONVERSION RATE: 2.5%

POWER: 80%

SIGNIFICANCE: 5%

TAILS: ONE

MDE: 10%

YOUR ORDER SUMMARY

YOUR ITEM	ITEM PRICE	QUANTITY	PRICE
John Varvatos Commuter Cap Toe Derby **SKU:** #8413525 **COLOR:** CHARCOAL **SIZE:** 12 **WIDTH:** D - MEDIUM	**$228.00** **ONLY 3 LEFT!** This item may become unavailable if someone else purchases it.	1 Remove	**$228.00**

SUBTOTAL (1 ITEM):	**$228.00**
NEXT DAY SHIPPING:	**FREE**
* TAX TO BE COLLECTED:	**$0.00**
STORE CREDIT FOR KYLE RUSH:	**-$45.99**
GRAND TOTAL:	**$182.01**

COMPLETE YOUR ORDER

By placing your order, you agree to Zappos.com's Privacy Policy and Terms of Use.

Shipping and Returns Policy

SUBMIT MY ORDER

CHECKOUT PAGE ASSUMPTIONS

1. Lower traffic
2. Relatively high conversion rate

Let's assume:

- 50% conversion rate
- 10,000 monthly unique visitors

RUN EXPERIMENT FOR 1 MONTH (10K VISITORS)

CONVERSION RATE: 50%

POWER: 80%

SIGNIFICANCE: 5%

TAILS: ONE

MDE: 5%

Start by focusing a/b tests on the last step in your conversion funnel.

WHAT SHOULD I TEST?

Depends on MDE and time.

LANDING PAGE OPTIMIZATION

(MOZCON 2013)

bit.ly/1wkpgye

GOING BEYOND THE LOW HANGING FRUIT

(CONVERSION CONFERENCE 2014)

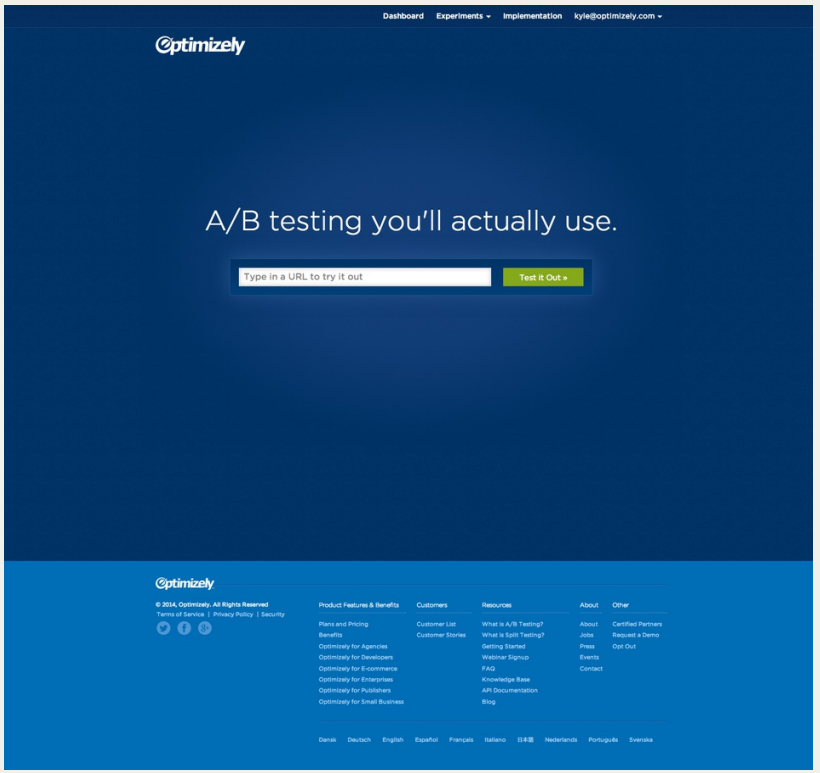
bit.ly/1kU4sZ0

BE FEARLESS

EXPERIMENT

Terse vs verbose

Optimizely
Features & Benefits
Pricing
Customers
Resources
Jobs
A/B testing you'll actually use.
Website optimization made easy.
Type in a URL to try it out
Test it Out »
REGISTRATION NOW OPEN FOR OPTICON 2014
Over 6,000 Happy Customers
FICO
FOX
the guardian
CBS
The Walt Disney Company
Track what matters to you
Say goodbye to technical bottlenecks and hello to actionable data
Revenue now, deploy later
100%
Try Optimizely FREE for 30 days!
Sign Up Today
In the News
The New York Times
Forbes
WIRED


Dashboard
Experiments
Implementation
kyle@optimizely.com
Optimizely
A/B testing you'll actually use.
Type in a URL to try it out
Test it Out »
© 2014, Optimizely. All Rights Reserved
Terms of Service | Privacy Policy | Security
Product Features & Benefits
Customers
Resources
About
Other
Plans and Pricing
Benefits
Optimizely for Agencies
Optimizely for Developers
Optimizely for E-commerce
Optimizely for Enterprises
Optimizely for Publishers
Optimizely for Small Business
Customer List
Customer Stories
What is A/B Testing?
What is Split Testing?
Getting Started
Webinar Signup
FAQ
Knowledge Base
API Documentation
Blog
About
Jobs
Press
Events
Contact
Certified Partners
Request a Demo
Opt Out
Dansk
Deutsch
English
Español
Français
Italiano
日本語
Nederlands
Português
Svenska

RESULT

+31% LEADS

99.9% CONFIDENCE

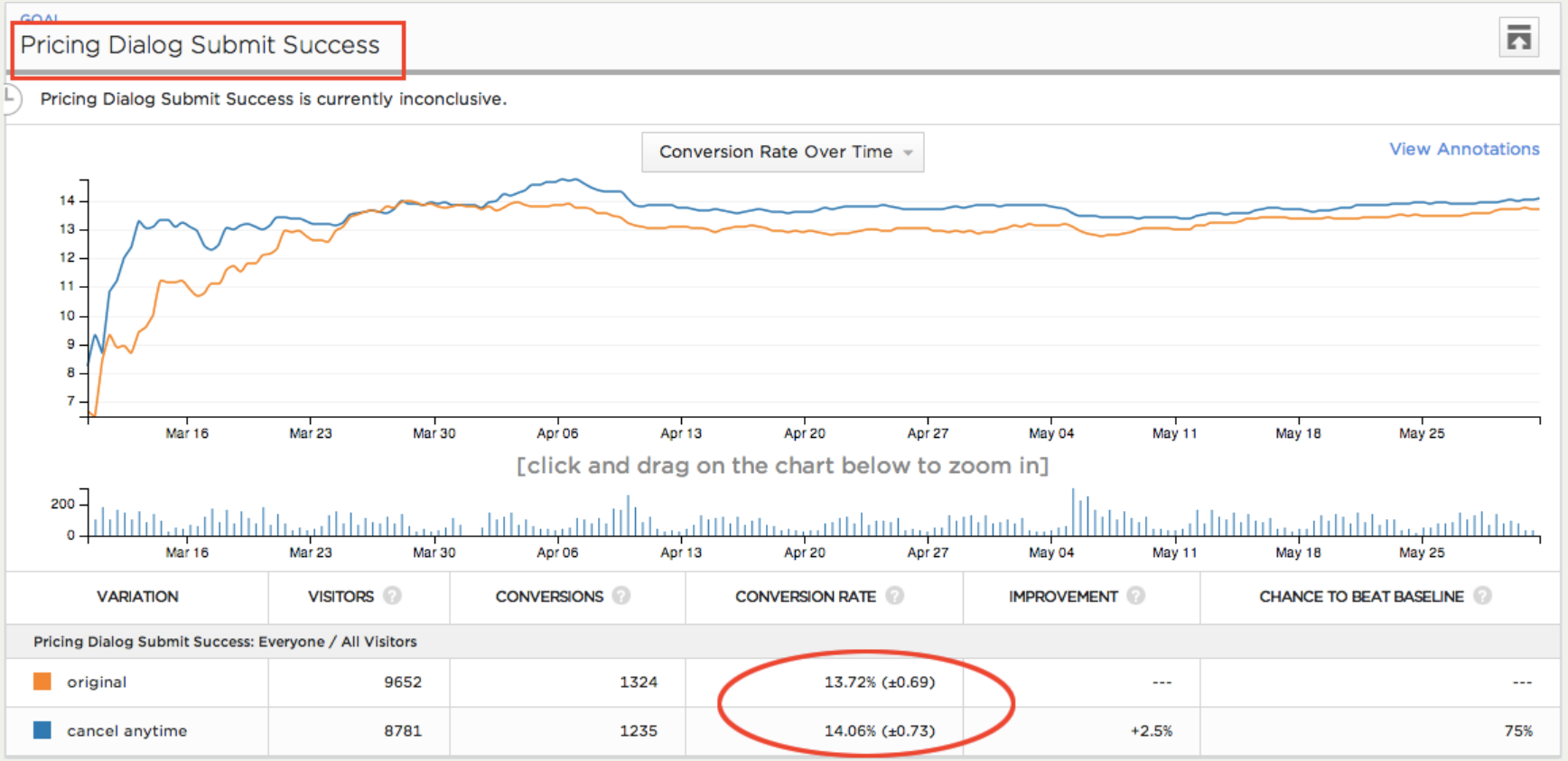
BEFORE YOU TEST

GOALS

- Measure as many goals as possible
  - micro: form field errors, time on page
  - macro: purchase, revenue
- Choose a primary goal
- Don't forget about down the funnel goals
  - Repeat purchase
  - Save payment information

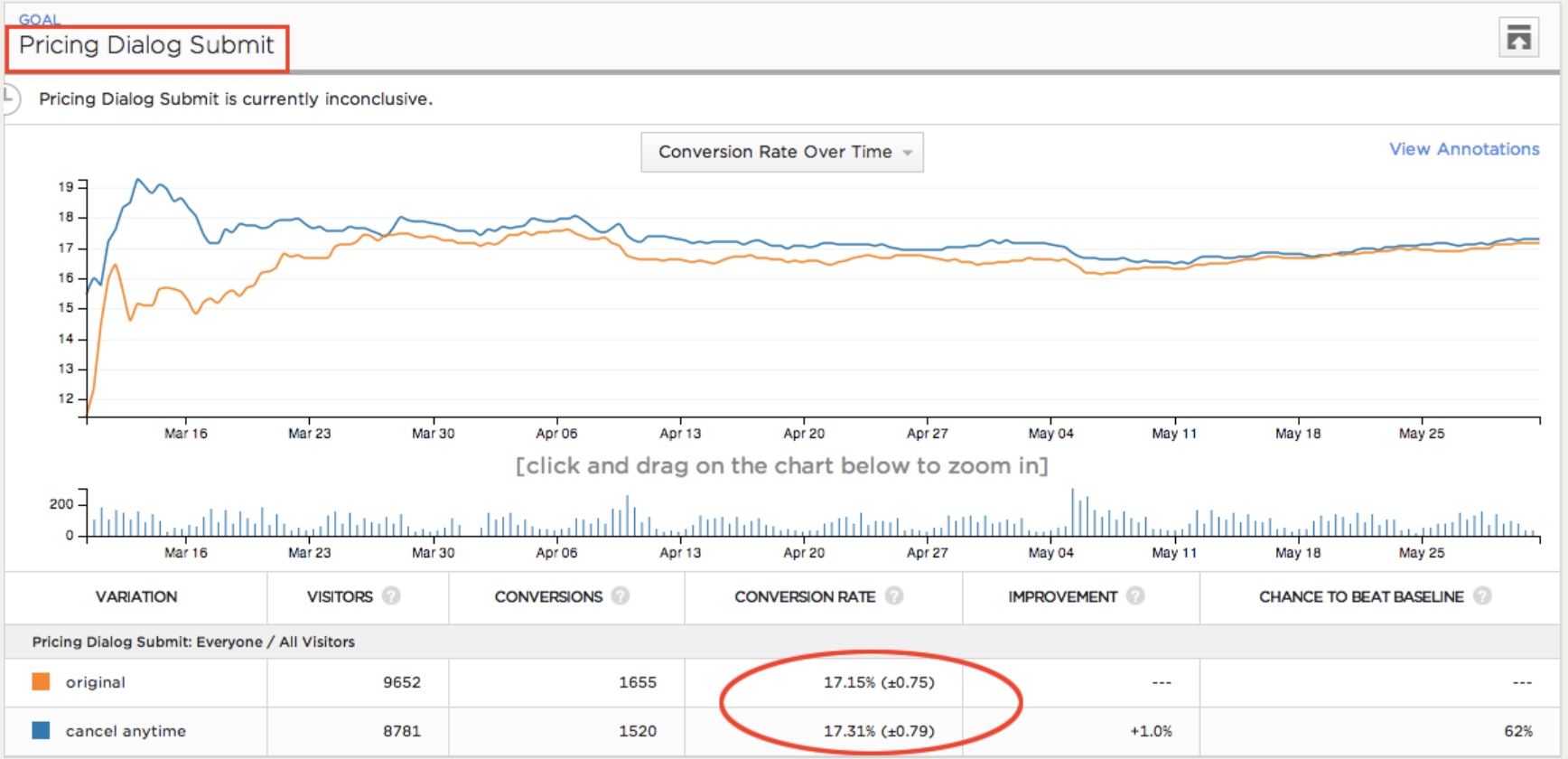
EXAMPLE

Success vs. submit



VARIATION	VISITORS	CONVERSIONS	CONVERSION RATE	IMPROVEMENT	CHANCE TO BEAT BASELINE
Pricing Dialog Submit Success: Everyone / All Visitors					
original	9652	1324	13.72% (±0.69)	---	---
cancel anytime	8781	1235	14.06% (±0.73)	+2.5%	75%

13.72% CONVERSION RATE



VARIATION	VISITORS	CONVERSIONS	CONVERSION RATE	IMPROVEMENT	CHANCE TO BEAT BASELINE
Pricing Dialog Submit: Everyone / All Visitors					
original	9652	1655	17.15% (±0.75)	---	---
cancel anytime	8781	1520	17.31% (±0.79)	+1.0%	62%

17.15% CONVERSION RATE

25% DIFFERENCE

Measure as many goals as possible for every experiment.

QUALITY ASSURANCE

- No bugs in the variation
- No bugs in the control
- Tracking works correctly

ELIMINATING BIAS

Double blind experiments

EXPERIMENT BRIEF

- Hypothesis
- Audience description
- Goals tracked
- Stopping conditions
- Screenshots
- QA summary

EXPERIMENT BRIEF TEMPLATE

bit.ly/1nvNJjx

AFTER YOU TEST

STATISTICAL TIE

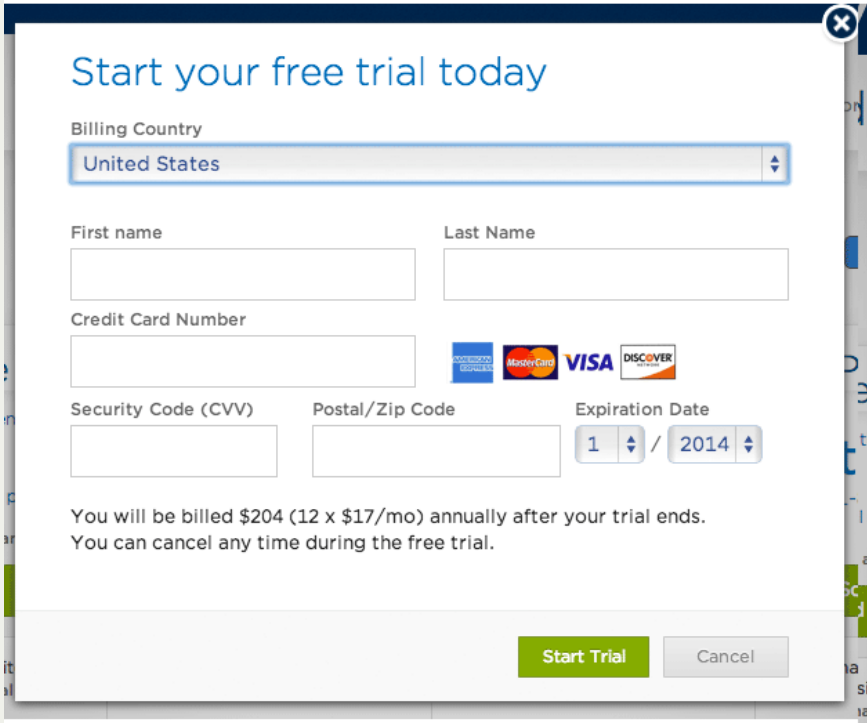
Not enough data to conclude that there is a difference.

The overwhelming majority of experiments results are a statistical tie.

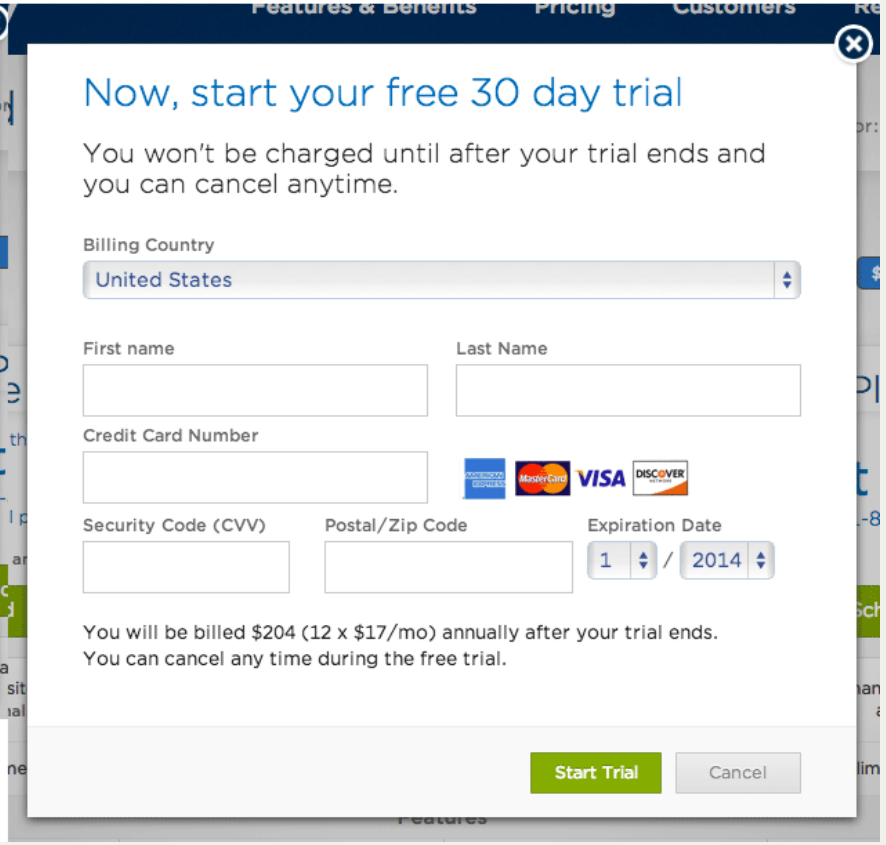
EXAMPLE

Statistical tie

control



variation



RESULT

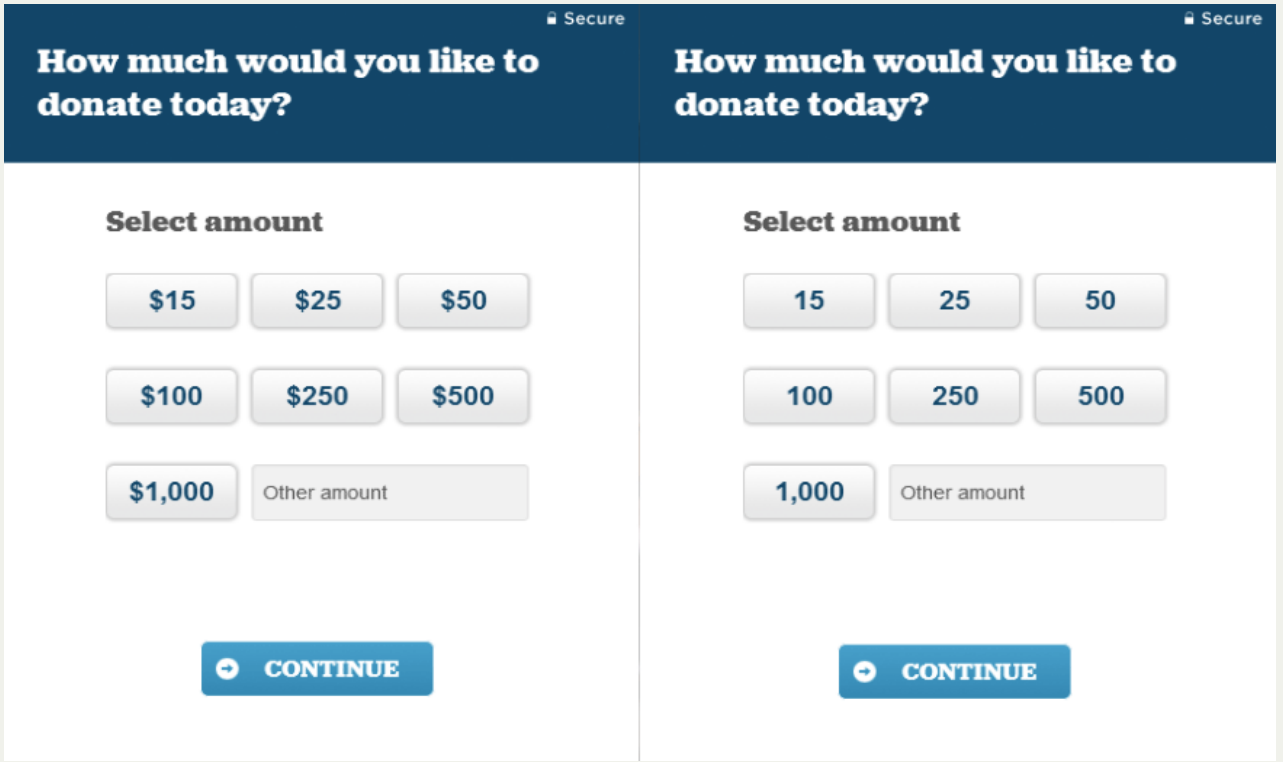
VARIATIONS	VISITORS	CONVERSIONS	CONVERSION RATE	IMPROVEMENT	CHANCE TO BEAT BASELINE
cancel anytime	10,239	1,469	14.3% (±0.68%)	+2.7%	78.4%
original BASELINE	11,165	1,560	14.0% (±0.64%)	---	---

RETESTING

EXAMPLE

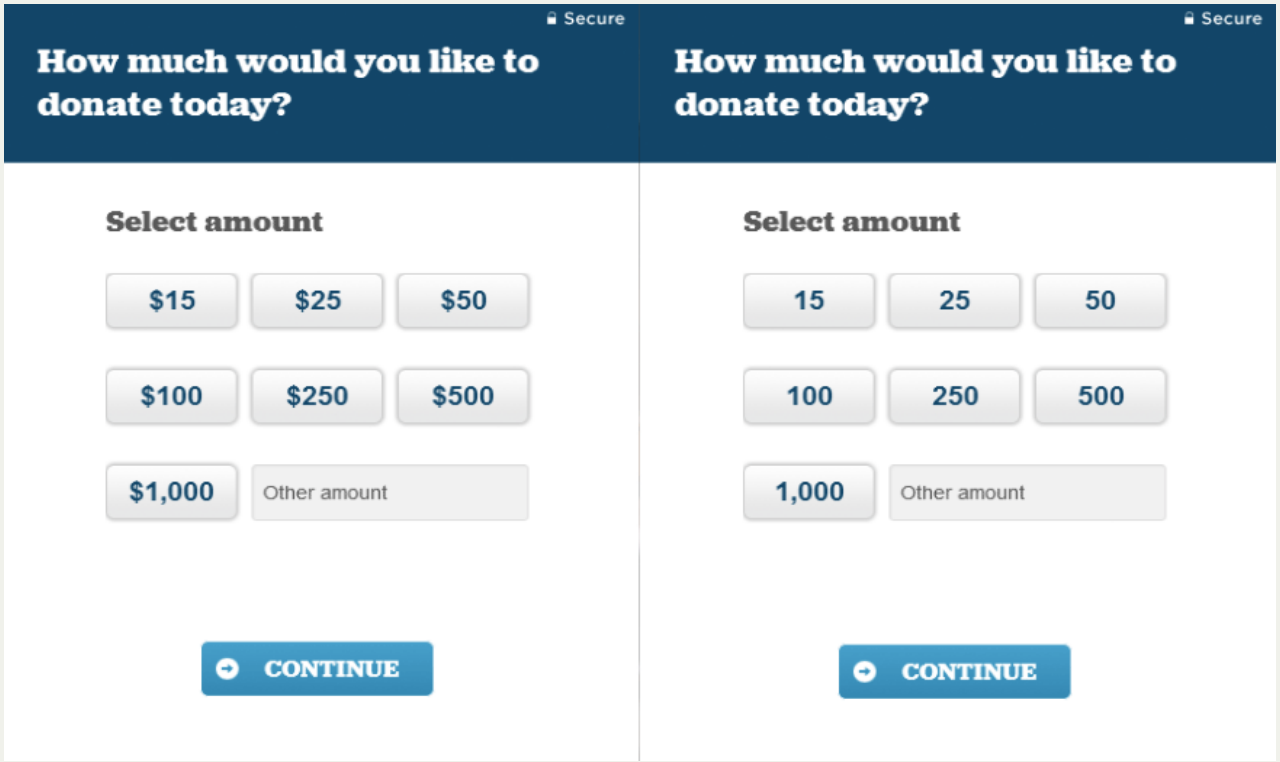
Retesting

FIRST TEST



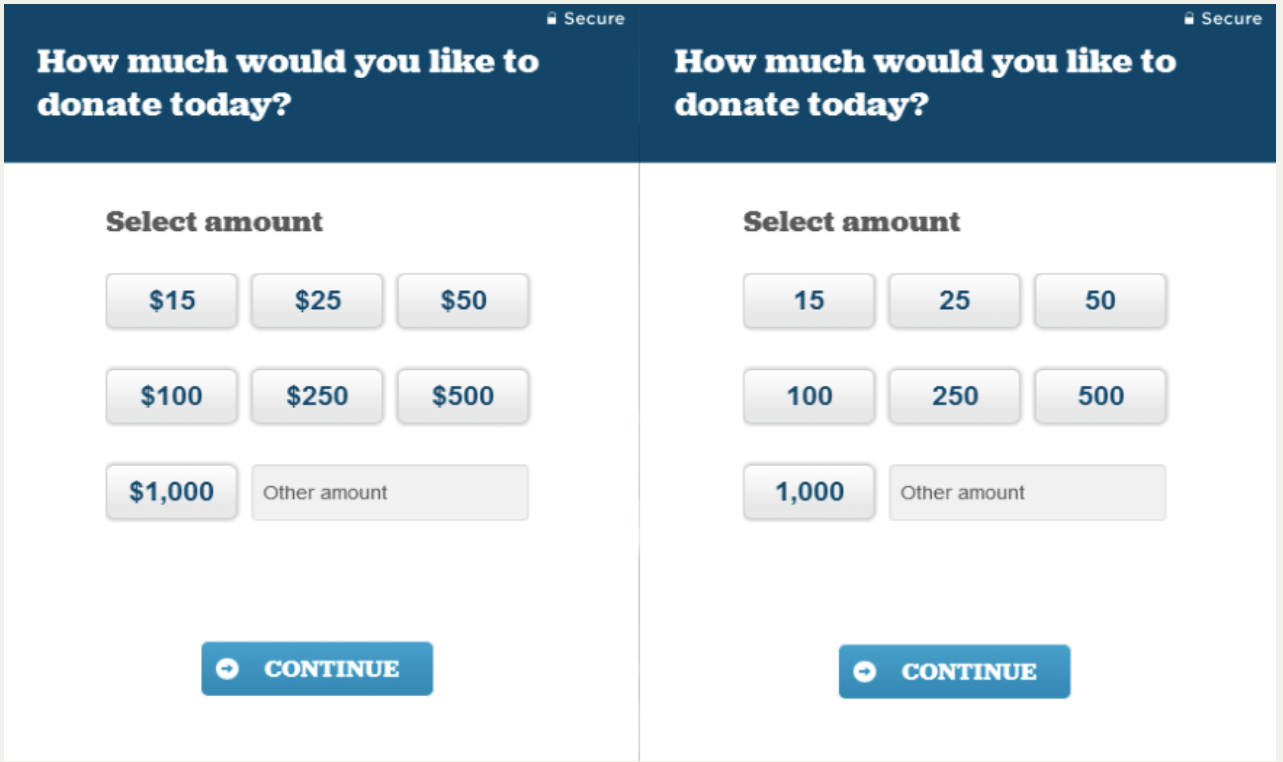
+36% revenue

SECOND TEST



statistical tie

THIRD TEST



Statistical tie

SHARE YOUR RESULTS

Always, always record detailed experiment results in an archive.

EXPERIMENT ARCHIVE

- experiment date
- audience/url
- screenshots
- hypothesis
- results
- link to experiment
- link to result csv

EXPERIMENT ARCHIVE TEMPLATE

bit.ly/1q9tRWI

MOZCON

2014

THANK YOU!

SESSION Q&A

Kyle Rush • Optimizely

@kylerush • kyle@optimizely.com

#MozCon