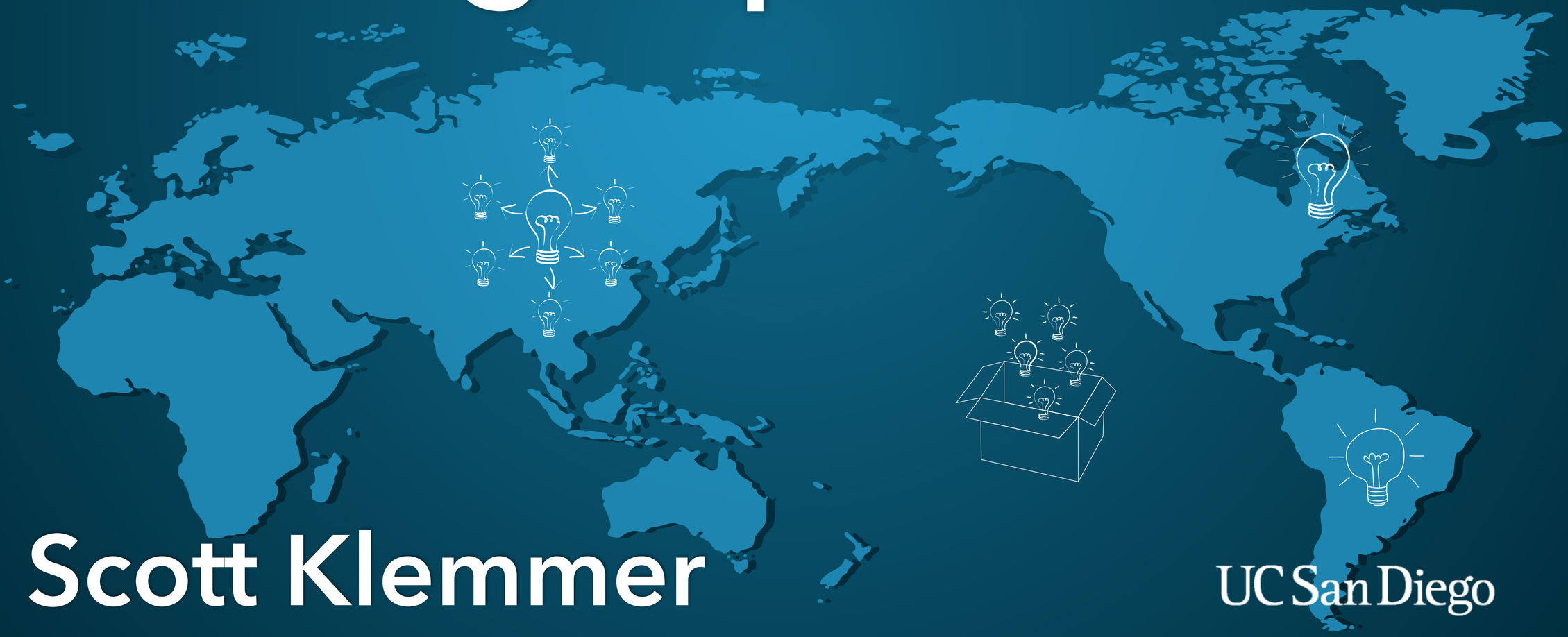
# Learning from & Running Experiments

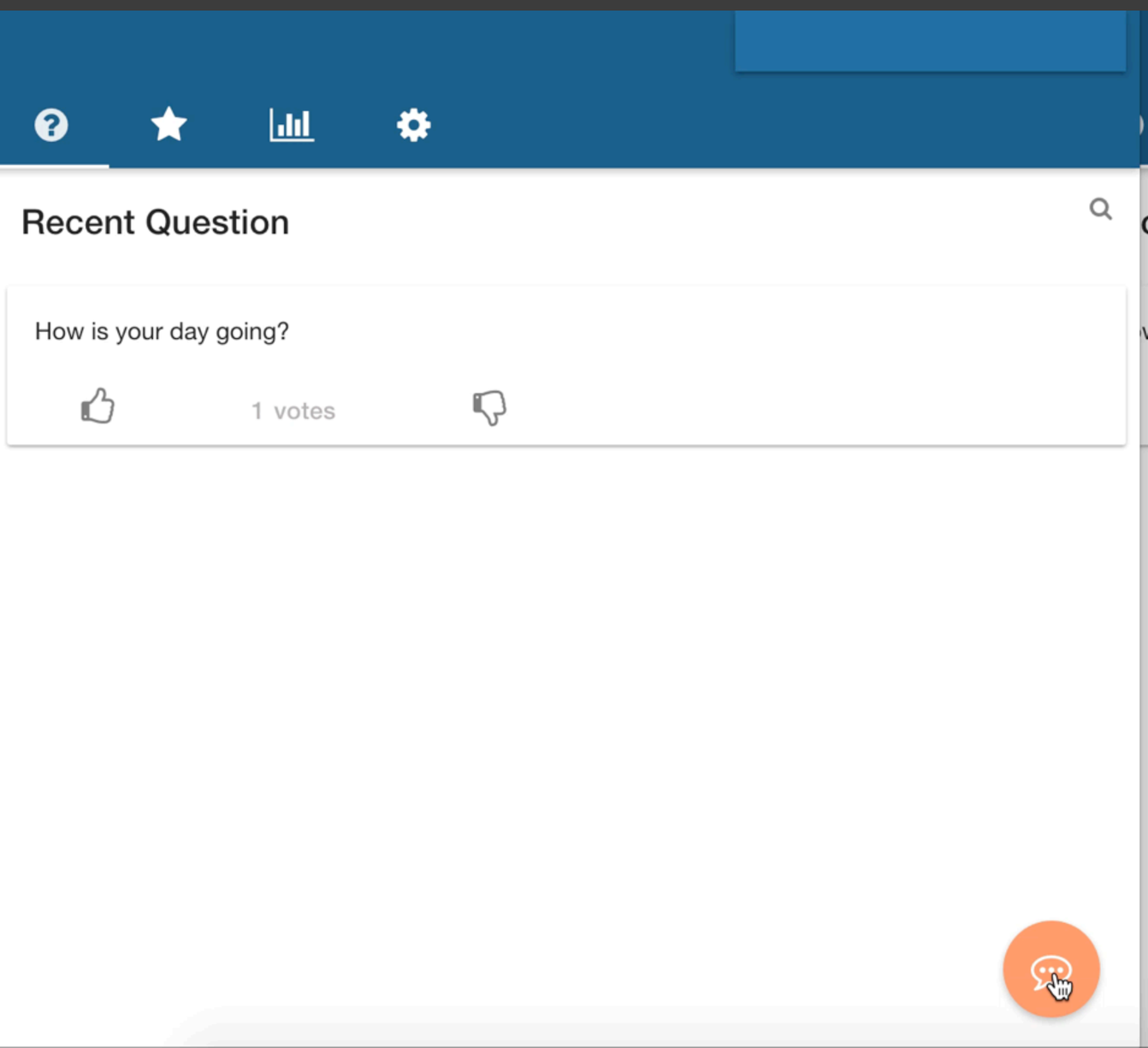


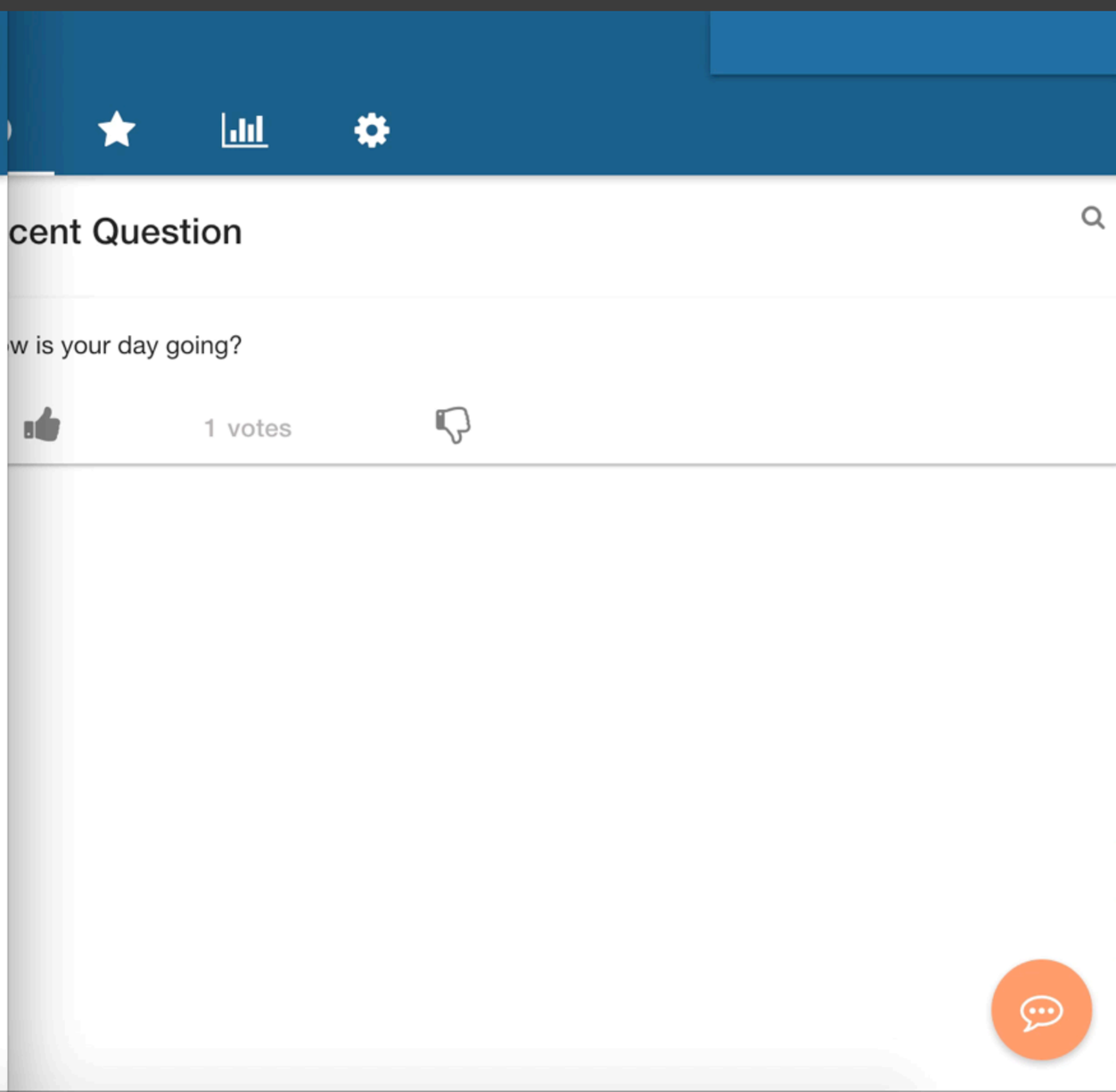
Scott Klemmer

UC San Diego

# A6 Example

# Kyly Vass, Austin Cheng, & Christopher Stringham



# “Do You Like My Interface?”

“How much do you like my interface?”

“This is a useful interface: agree/disagree”

# Please the Experimenter Bias

# Developers are Valuable Testers

# Getting beyond "*do you like my interface?*"

- What's the comparison?
- What's the yardstick?

# Getting beyond "*do you like my interface?*"

- **Baserates**: How often does Y occur?
  - Requires measuring Y.

# Getting beyond "*do you like my interface?*"

- Baserates: How often does Y occur?
  - Requires measuring Y.

- **Correlations**: Do X and Y co-vary?
  - Requires measuring X and Y.

# Getting beyond "*do you like my interface?*"

- Baserates: How often does Y occur?
  - Requires measuring Y.

- Correlations: Do X and Y co-vary?
  - Requires measuring X and Y.

- **Causes**: Does X cause Y?
  - Requires measuring X and Y, and manipulating X.
  - Also requires somehow accounting for the effects of other independent variables (confounds)!

# Let’s introduce a few terms...

manipulations

# **Independent** Variables

measures

# **Dependent** Variables

precision

# **Internal** Validity

generalizability

# **External** Validity

# IS MY COOL NEW APPROACH BETTER THAN THE INDUSTRY STANDARD?

Friday, August 17, 2007 11:03 AM PT Posted by Harry McCracken

# A Not-Very-Useful iPhone Keyboard Study

ADD TO MY PAGES PRINT E-MAIL COMMENT RSS

SLASHDOT IT DIGG THIS DEL.ICIO.US NEWSVINE

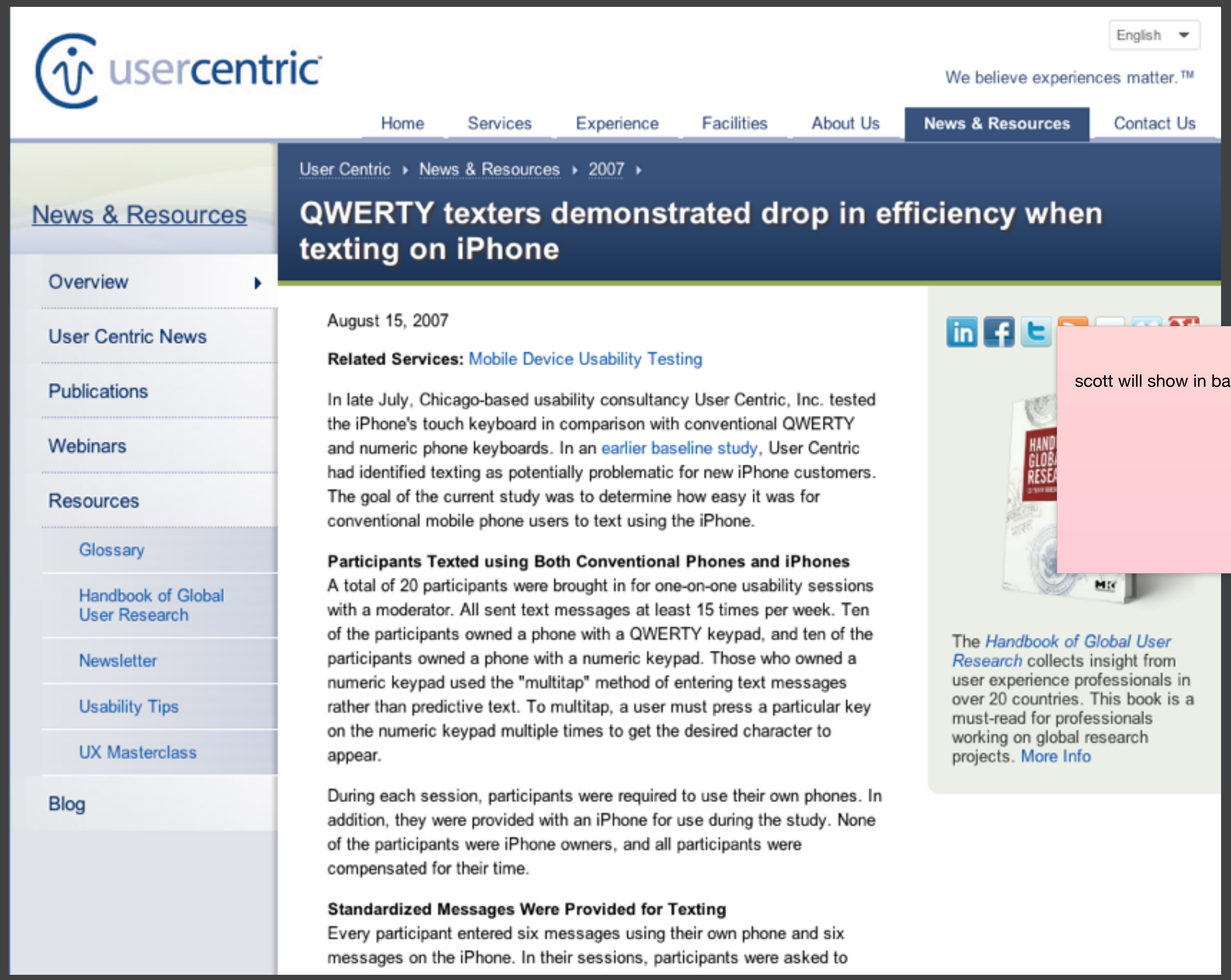
Research firm User Centric has released a study that tries to gauge how effective the iPhone's unusual on-screen keyboard is. The goal is certainly a noble one, but I can't say that the survey's approach results in data that makes much sense.

User Centric brought in twenty owners of other phones--half who had ones with QWERTY keyboards, and half who had ordinary numeric phone keypads. None were familiar with the iPhone. The research involved having the test subjects enter six sample text messages with the phones they already had, and six with an iPhone.

Logical end result: These iPhone newbies took twice as long to enter text with an iPhone as they did with their own phones, and made lots more typos.

Source: PC World

# Benefits and Drawbacks?




scott will show in background

- Manipulation: Input Style
- Measure: Words per minute
- External validity: not so much

# A better version: actual users

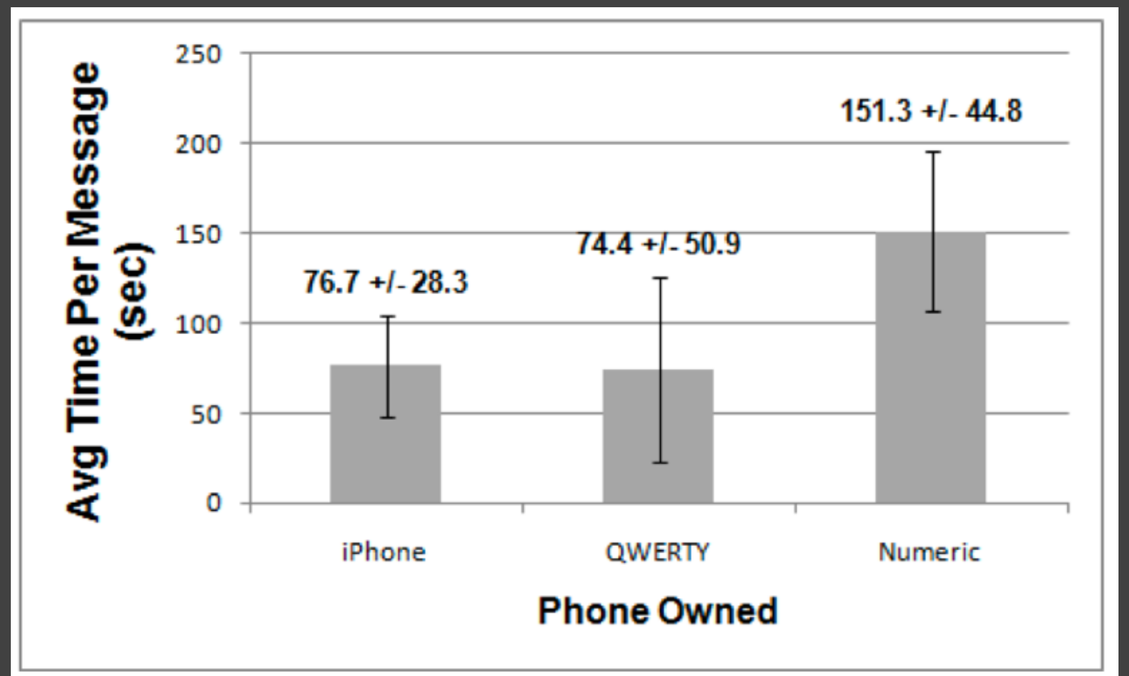



Figure 1. Average time to type a message on phones owned by the participants (*M* +/- *SD*).

- Manipulation: Input Style
- Measure: Words per minute
- ...and error rate

An In-Depth Look into the Text Entry User Experience on the iPhone, Jennifer M. Allen, Leslie A. McFarlin and Thomas Green, Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2008

# iPhone & Qwerty users similar speed, but make more errors

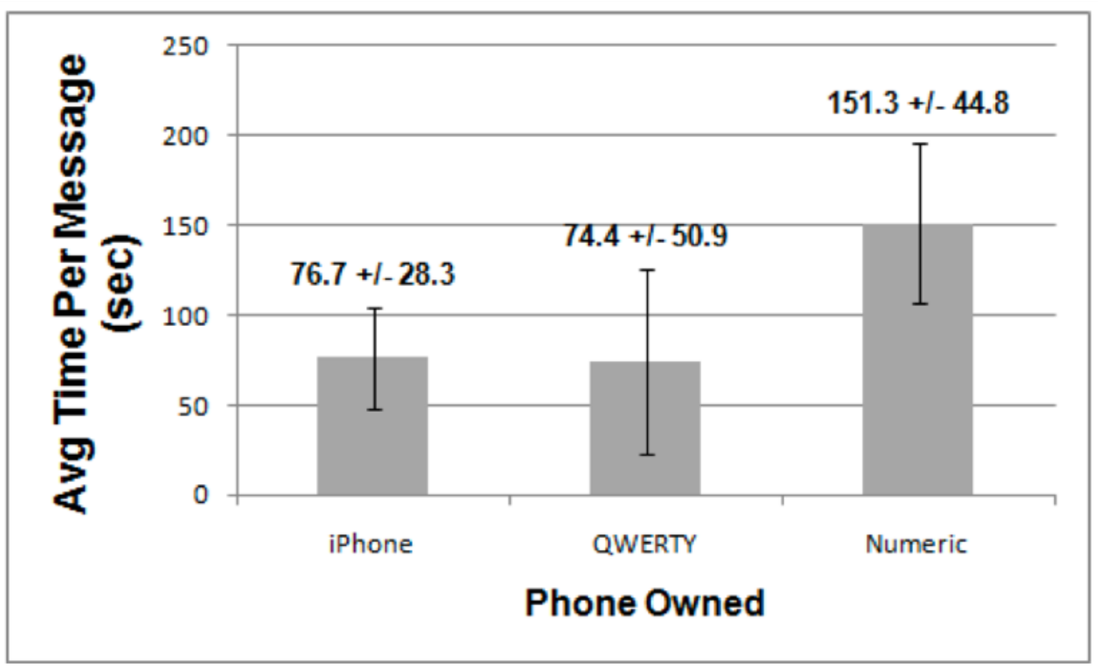



Figure 1. Average time to type a message on phones owned by the participants (*M* +/- *SD*).

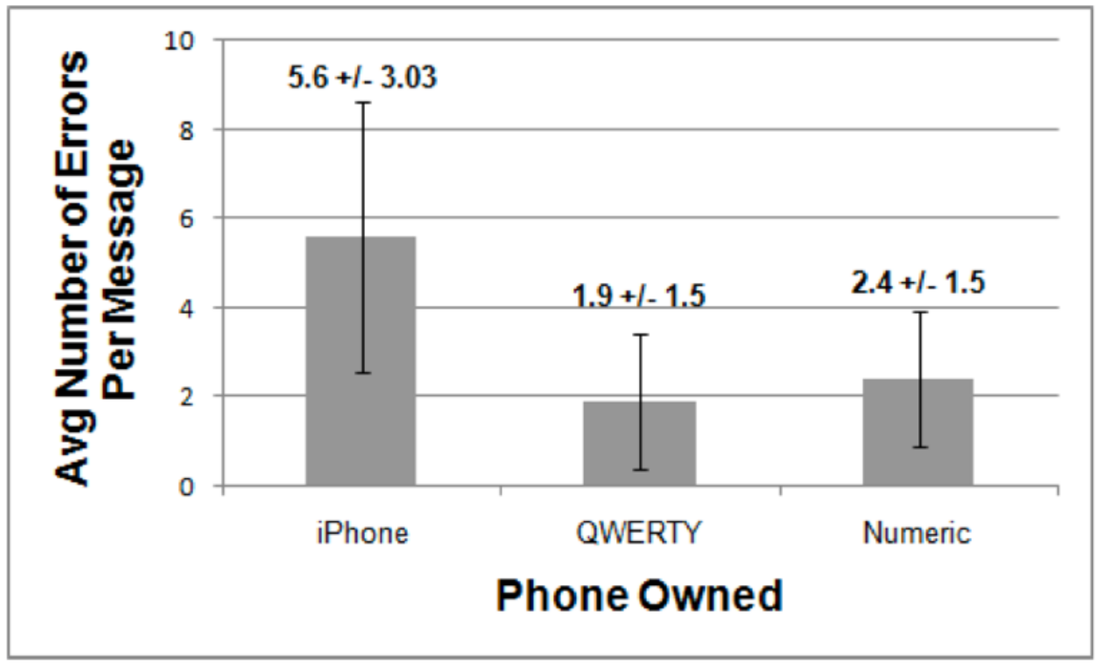



Figure 2. Average number of total errors per message made by participants using their own phones (*M* +/- *SD*).

An In-Depth Look into the Text Entry User Experience on the iPhone, Jennifer M. Allen, Leslie A. McFarlin and Thomas Green, Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2008

# Strategies for fairer comparisons

- Insert your new approach into the production setting
- Recreate the production approach in your new setting
- Scale things down so you're just looking at a piece of a larger system
- When expertise is relevant, train people up

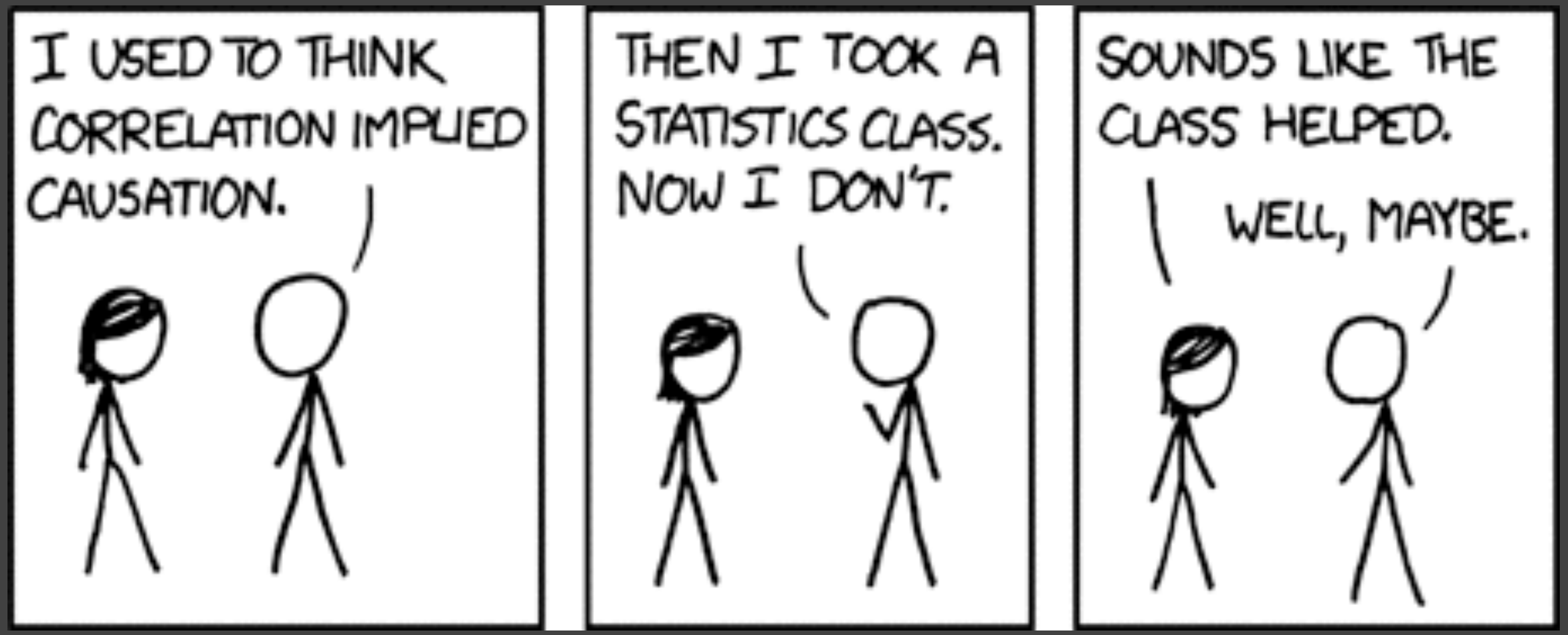
# IS INTERFACE X BETTER THAN INTERFACE Y?

Most often, the answer is
“it depends”

The question is “on what”?

# CONTROLLED COMPARISON ENABLES CAUSAL INFERENCE

*(a fancy way of saying you can learn stuff)*




http://xkcd.com/552

# Make a Study Plan

- Who are your users?
- What's your task?
- What will you measure?

# Turn to the person next to you

- In A7, you'll submit a 2-3 sentence plan for the study you'll run in A8
- What do you hope to learn? What questions do you have about your Webapp?
- Help each other improve your plan to better match what you want to learn

# Running Web Experiments

# National Alert Registry

*Be Aware • Be Alert • Be Safe*

## Search Results

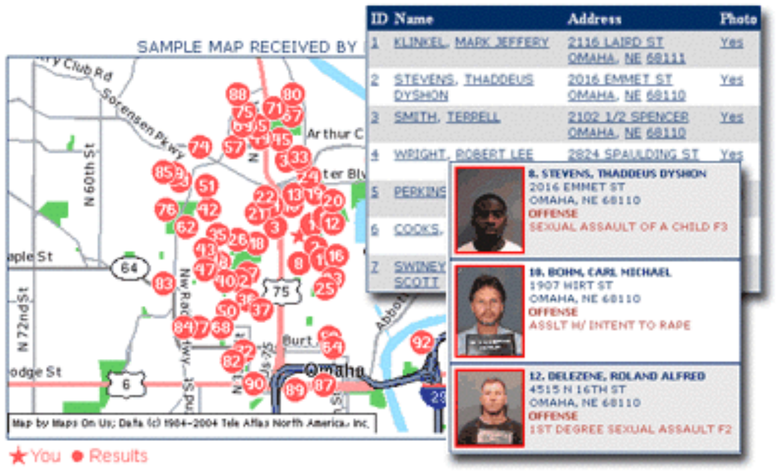
| | | | |
|---|---|---|---|
| **Search Initiated:** | 8/2/2005 | **Time:** 12:12 P.M. | |
| **Search Location:** | **Zip:** 32225 | **City:** JACKSONVILLE, JAX | **State:** FL |

**33 Registered sex offenders are located in your Immediate area.**

Get the **complete report** including map, **photograph**, and addresses to find out who they are and where they **live**. You will be able to enter your full address and specify a **larger** radius.

**The complete report includes:**

- Photos
- Appearance Details
- Conviction Information
- Offense Dates
- Offense Details
- Street Addresses
- Name Aliases




See the entire sample Report!

To get a complete Predator Report **today**, register below with National Alert Registry, and for just $10.00 we will create a custom report using your address. We will also send you a "Red Alert" notification when new predators are in your area.

Email:

(Email address used for RED ALERT Notifications)

First Name:

Last Name:

Address 1:

Address 2:

City:

### Background

The **U.S. Congress** passed three laws that require States to keep **track** of sex offenders; the Wetterling Registration Act, the Lynchner Tracking and **Identification Act** and Megan's Law. On March 5, 2003, The Supreme Court ruled that information about **potential predators** may be posted on the Internet.

### The Problem

The information is available and you have the **right** to see it. But the problem is that access is very **difficult**. You have to know the name of the individual you are looking for and in many States you must go to your police station and complete a **request form**. In some States there is a fee and **limit** to view only two names

### The Solution - Predator Report

Now in the **privacy** of your home you can view the same information displayed on a map of **your neighborhood**. The free service provides the number of predators in your neighborhood. You can then choose to purchase the full service, which includes details about those **sexual**

# National Alert Registry

*Be Aware • Be Alert • Be Safe*

## Search Results for Zip Code 32225

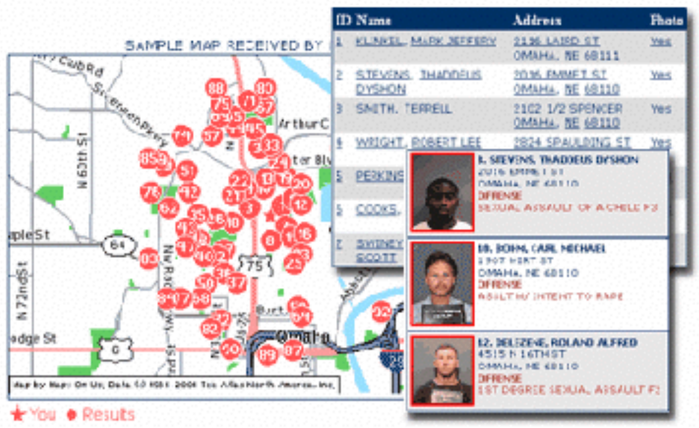
| | | | |
|---|---|---|---|
| **Search Initiated:** | 8/2/2005 | **Time:** 12:11 P.M. | |
| **Search Location:** | **State:** FL | **City:** JACKSONVILLE, JAX | **Zip:** 32225 |
| There are currently 32 Registered Sex Offenders located in your zip code | | | |

Would you like **names**, **pictures** and **addresses**? Click Here

The National Alert Registry has made all this vital information available to you. Plus you can get a custum map of your neighborhood, identifying your home and where offenders reside in proximity to it.

**Your complete "Predator Report" includes all of the below:**

- Color Photos (when available)
- Appearance Details
- Street Addresses
- A Detailed Map
- Name & Aliases
- Conviction Information
- Offense Dates
- Offense Details




See the entire sample Report!

Your can purchase your complete Predator Report and **access this vital information today**. For just $10.00 we will create a custom report using your address.

Email:
(Email address used for RED ALERT Notifications)
First Name:
Last Name:
Address 1:
Address 2:
City:
State: Select
Zip: 32225
Country: United States
Phone:

**Billing Information**

By submitting this information you agree to a one time

*"I love my NAR monthly Predator Report. You can't imagine my shock to find a registered sex offender living one mile from my home. I have grandchildren that visit often and you can bet I will keep a MUCH closer watch on them from now on. NAR is a wonderful tool for keeping my family safe." Jolene S., Florida*

*"NAR Predator Report is wonderful! I live in an upscale neighborhood and thought I was safe, but when I checked my Predator Report I discovered that we have three (3) offenders living in our neighborhood. I was astounded."
Bill S., New York*

*"I am very pleased that I subscribed to this. It does indeed, keep me:
ALERT, and AWARE and I feel SAFER! THANK YOU for these services."
-- Justice of Colorado*

VeriSign

# National Alert Registry

*Be Aware • Be Alert • Be Safe*

## Free Search Results for Zip Code 32225

| | | | |
|---|---|---|---|
| **Search Initiated:** | 8/2/2005 | **Time:** 12:12 P.M. | |
| **Search Location:** | **State:** FL | **City:** JACKSONVILLE, JAX | **Zip:** 32225 |
| 32 Registered Sex Offenders currently located in your zip code | | | |

## Create Your Sex Offender Map And Detailed Predator Report

Get a custom **map of your neighborhood**, identifying your home and where offenders reside in your area. It only takes a few seconds to generate your report and costs **just $10**.

### Your Neighborhood Report

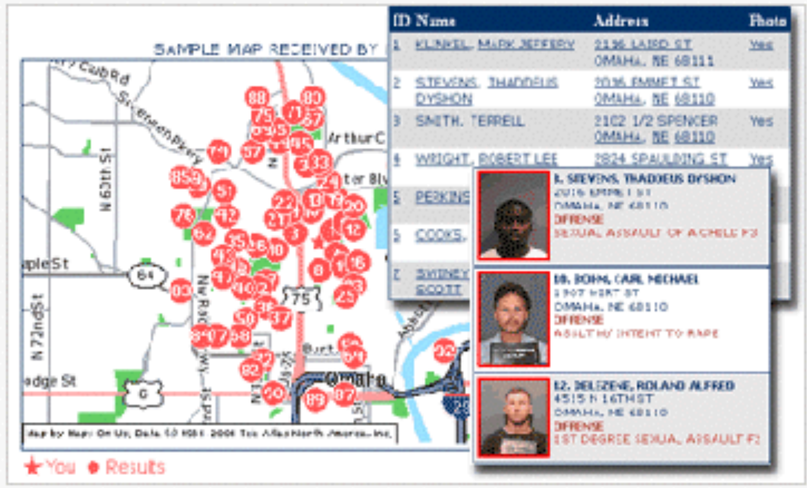
We will use your address to create a custom report

Email:
(Only used for RED ALERT Notifications)
First Name:
Last Name:
Address 1:
Address 2:
City:
State: Select
Zip: 32225
Country: United States
Phone:

**Select your billing preference:**

Single "Predator Report" one time charge of **$10.00**

### Sample Report




See the entire sample Report!

**The complete Predator Report includes:**

- **Color Photos** (when available)

# A/B/C Split Test

| | PAGE A | PAGE B | PAGE C |
|---|---|---|---|
| Percent of traffic | 34% | 33% | 33% |
| New sales | 244 | 282 | 114 |
| Change | N/A | 15.57% | -53.28% |

NATIONAL ALERT REGISTRY *Be Aware • Be Alert • Be Safe*

**Search Results**

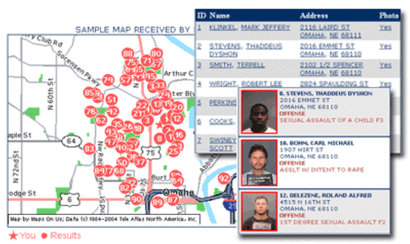
| | | | |
|---|---|---|---|
| **Search Initiated:** | 8/2/2005 | **Time:** 12:12 P.M. | |
| **Search Location:** | **Zip:** 32225 | **City:** JACKSONVILLE, JAX | **State:** FL |

33 Registered sex offenders are located in your Immediate area.

Get the **complete report** including map, **photograph**, and addresses to find out who they are and where they **live**. You will be able to enter your full address and specify a **larger** radius.

**The complete report includes:**

- Photos
- Appearance Details
- Conviction Information
- Offense Dates
- Offense Details
- Street Addresses
- Name Aliases




See the entire sample Report!

To get a complete Predator Report **today**, register below with National Alert Registry, and for just $10.00 we will create a custom report using your address. We will also send you a "Red Alert" notification when new predators are in your area.

Email:
(Email address used for RED ALERT Notifications)
First Name:
Last Name:
Address 1:
Address 2:
City:
State: Select
Zip: 32225
Country: United States
Phone:

**Billing Information**

**Background**

The **U.S. Congress** passed three laws th require States to kee **track** of sex offende the Wetterling Registration Act, the Lynchner Tracking a **Identification Act** Megan's Law. On Ma 2003, The Supreme ruled that informatio about **potential predators** may be posted on the Intern

**The Problem**

The information is available and you ha **the right** to see it. B the problem is that a is very **difficult**. You have to know the na the individual you ar looking for and in ma States you must go t your police station a complete a **request form**. In some State there is a fee and **lir** view only two names

**The Solution - Predator Report**

Now in the **privacy** your home you can the same informatior displayed on a map **your neighborhoo** The free service pro the number of preda in your neighborhoo You can then choose purchase the full ser which includes detail about those **sexual offenders**, including specific addresses. T full service search is based on **your** exac address.

NATIONAL ALERT REGISTRY *Be Aware • Be Alert • Be Safe*

**Search Results for Zip Code 32225**

| | | | |
|---|---|---|---|
| **Search Initiated:** | 8/2/2005 | **Time:** 12:11 P.M. | |
| **Search Location:** | **State:** FL | **City:** JACKSONVILLE, JAX | **Zip:** 32225 |

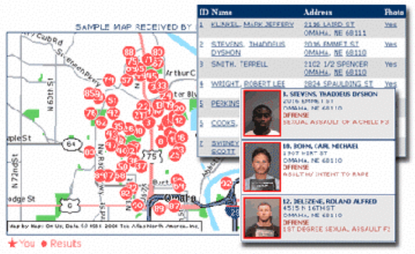
There are currently 32 Registered Sex Offenders located in your zip code

Would you like **names**, **pictures** and **addresses**? Click Here

The National Alert Registry has made all this vital information available to you. Plus you can get a custum map of your neighborhood, identifying your home and where offenders reside in proximity to it.

**Your complete "Predator Report" includes all of the below:**

- Color Photos (when available)
- Appearance Details
- Street Addresses
- A Detailed Map
- Name & Aliases
- Conviction Information
- Offense Dates
- Offense Details




See the entire sample Report!

Your can purchase your complete Predator Report and **access this vital information today**. For just $10.00 we will create a custom report using your address.

Email:
(Email address used for RED ALERT Notifications)
First Name:
Last Name:
Address 1:
Address 2:
City:
State: Select
Zip: 32225
Country: United States
Phone:

**Billing Information**

By submitting this information you agree to a one time

*"I love my NAR monthly Predator Report. You can't imagine my shock to find registered sex offender liv one mile from my home. I grandchildren that visit o and you can bet I will kee MUCH closer watch on the from now on. NAR is a wonderful tool for keeping family safe." Jolene S., Flo*

*"NAR Predator Report is wonderful! I live in an up neighborhood and thought was safe, but when I chec my Predator Report I discovered that we have t (3) offenders living in our neighborhood. I was astounded." Bill S., New York*

*"I am very pleased that I subscribed to this. It doe indeed, keep me: ALERT, and AWARE and I f SAFER! THANK YOU for the services." -- Justice of Colorado*

VeriSign

NATIONAL ALERT REGISTRY *Be Aware • Be Alert • Be Safe*

**Free Search Results for Zip Code 32225**

| | | | |
|---|---|---|---|
| **Search Initiated:** | 8/2/2005 | **Time:** 12:12 P.M. | |
| **Search Location:** | **State:** FL | **City:** JACKSONVILLE, JAX | **Zip:** 32225 |

32 Registered Sex Offenders currently located in your zip code

**Create Your Sex Offender Map And Detailed Predator Report**

Get a custom **map of your neighborhood**, identifying your home and where offenders reside in your area. It only takes a few seconds to generate your report and costs **just $10**.

**Your Neighborhood Report**

We will use your address to create a custom report

Email:
(Only used for RED ALERT Notifications)
First Name:
Last Name:
Address 1:
Address 2:
City:
State: Select
Zip: 32225
Country: United States
Phone:

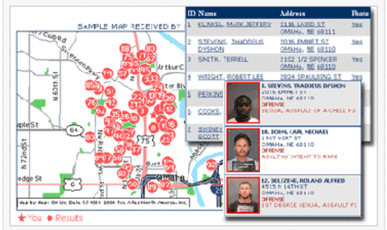
**Select your billing preference:**

- ○ Single "Predator Report" one time charge of **$10.00**
- ◉ Upgrade my registration to "**Red Alert**." Only **$4.95**/month more.

**Red Alert** includes:

1. **Automatically Updated Predator Reports** - Your map is updated daily when new information is available.
2. **Automatic Predator Notifications** when new offenders are in your area.
3. **Email alerts** when our nation-wide database has been updated.

**Sample Report**




See the entire sample Report!

**The complete Predator Report includes:**

- Color Photos (when available)
- Appearance Details
- Street Addresses
- A Detailed Map
- Name & Aliases
- Conviction Information
- Offense Dates
- Offense Details

# Ways design makes a difference

- Position and color of a call to action
- Position on the page of testimonials, if used
- Whether linked elements are in text or as images
- Amount of white space on a page, giving the content space to “breathe”
- Position and prominence of the main heading
- Number of columns used on the page
- Number of visual elements competing for attention
- Attributes of people and objects in photos

Content courtesy of A List Apart: http://www.alistapart.com/articles/designcancripple

# EDUCATORS CORNER
Entrepreneurship Education Resources

Brought to you by the
Stanford Technology Ventures Program

## Quick Shots

Spend just two minutes and a few odd seconds to learn bright ideas to build, grow, or sustain your entrepreneurial drive. These bite-sized, informative videos offer a fast jolt to your fast-tracking career.

Negotiating with Customers and Clients

Stan Christensen | Dan Springer

2 min. 1 sec.

Don't Kill Projects; Morph Them

Marissa Mayer

2 min. 2 sec.

Taking Risks

Vinod Khosla

2 min. 26 sec.

## New Podcasts

From medical devices to social responsibility, corporate acquisitions to personal tales of trial and triumph, our speakers share their collective wisdom and decades of business savvy.

Podcast: Representing the Socially Responsible Enterprise

Debra Dunn, Skoll Foundation | Jay Coen Gilbert, B Corporation | Bart Houlahan, B Corporation | Andrew Kassoy, B Corporation - 50 min. 56 sec.

The co-founders of B Lab, Jay Coen Gilbert, Bart Houlahan, and Andrew Kassoy, unveil their infrastructure play that seeks to give voice to the burgeoning panoply of green business. They explain how a higher set of corporate standards accountable to the environment, employees, and the community, can craft a healthier corporate

## Quick Shots

Spend just two minutes and a few odd seconds to learn bright ideas to build, grow, or sustain your entrepreneurial drive. These bite-sized, informative videos offer a fast jolt to your fast-tracking career.

Video: Negotiating with Customers and Clients

Stan Christensen, Stanford Technology Ventures Program; Arbor Advisors | Dan Springer, Responsys - 2 min. 1 sec.
Dan Springer, CEO of Responsys, acknowledges that the most frequent and important negotiations often occur with customers. Furthermore, Springer advises that the keys to a successful negotiation with customers include: 1) Preparation to discover what the customer wants and 2) Identifying the few things on which you will stand firm.

Video: Don't Kill Projects; Morph Them

Marissa Mayer, Google - 2 min. 2 sec.
Repackage, rejuvenate, re-market, and re-examine those products or practices you thought would fly, and craft them a new set of wings. Head of Google product development Marissa Mayer lives by the old adage that if at first you don't succeed, try again. She pushes aspiring business thinkers to breathe new life into failed ventures, as opposed to cutting the cord.

Video: Taking Risks

Vinod Khosla, Kleiner, Perkins, Caufield & Byers - 2 min. 26 sec.
Launching a start-up is not a rational act. And Vinod Khosla, a partner in Kleiner, Perkins, Caufield & Byers and former Sun Microsystems CEO, believes that success only comes from those who are foolish enough to think

# Results

- Version A (traditional version) was sent to 6272 users.
  Opened: 1638 - Click thrus: 722 - Forwards: 4
- Version B (Quick Shots version) was sent to 6263 users.
  Opened: 1769 - Click thrus: 922 - Forwards: 14

Courtesy Forrest Glick, Stanford STVP http://project8180.stanford.edu/category/wireframes/

GET INVOLVED
JOIN THE MOVEMENT
Email Address
Zip Code
SIGN UP
Media
Button
Baseline sign up rate: 8.26%
STANFORD

# Which option performed best?

- Sign up
- Learn More
- Sign up Now
- Join Us Now

**Courtesy Dan Siroker, Stanford HCI Seminar** http://hci.stanford.edu/courses/cs547/speaker.php?date=2009-05-08

# Now for the visual material: five options

Courtesy Dan Siroker, Stanford HCI Seminar http://hci.stanford.edu/courses/cs547/speaker.php?date=2009-05-08

Button: "Sign Up Now"

- Family Image
- Change Image
- Barack Video
- Springfield Video
- Sam's Video

Courtesy Dan Siroker, Stanford HCI Seminar http://hci.stanford.edu/courses/cs547/speaker.php?date=2009-05-08

STANFORD
UNIVERSITY

# Here We Saw

- Small changes: big difference
- Our expectations are often wrong

For Dustin Curtis, Which performed best?

- I'm on Twitter
- Follow me on twitter
- You should follow me on twitter
- You should follow me on twitter here

Courtesy Dustin Curtis: http://dustincurtis.com/you_should_follow_me_on_twitter.html

For Dustin Curtis, Which performed best?

**4.70%** I'm on Twitter

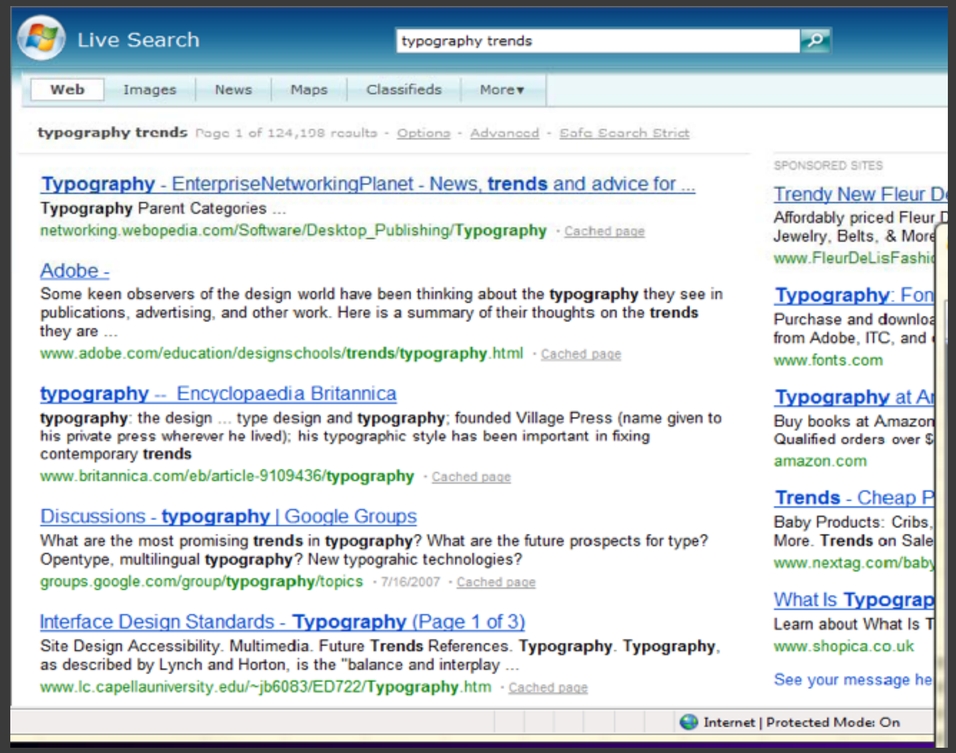
**7.31%** Follow me on twitter

**10.09%** You should follow me on twitter

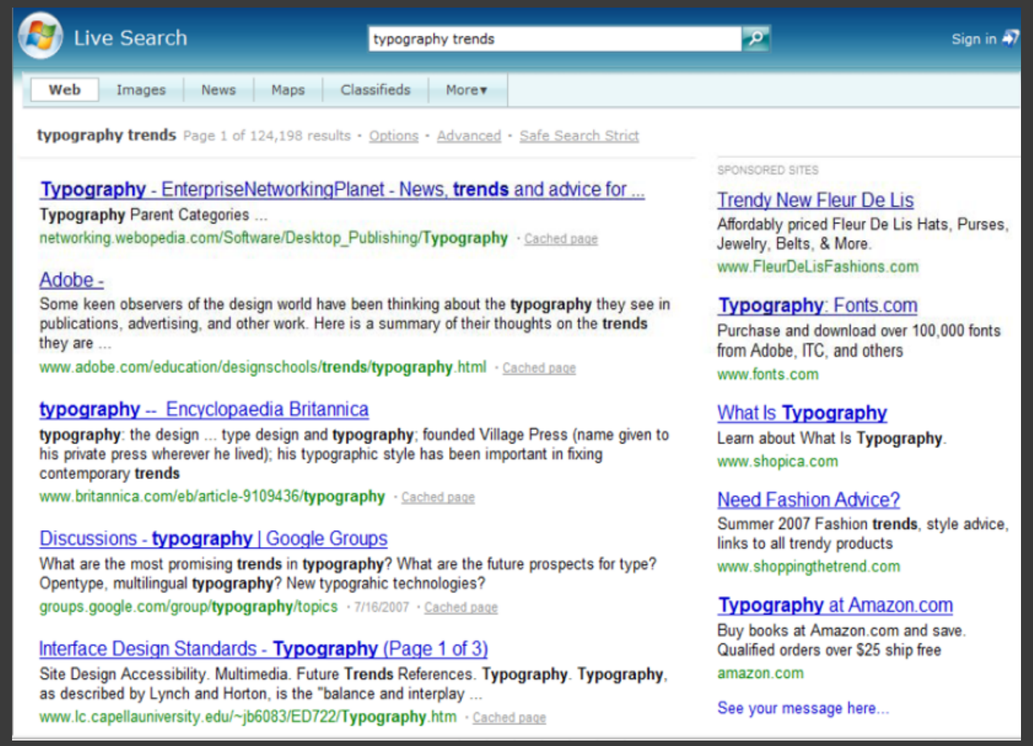
**12.81%** You should follow me on twitter here

Courtesy Dustin Curtis: http://dustincurtis.com/you_should_follow_me_on_twitter.html

# Typography Experiment:
## Color Contrast on MSN Live Search




A: Softer colors




B: High contrast

**B: Queries/User up 0.9%**

**Ad clicks/user up 3.1%**

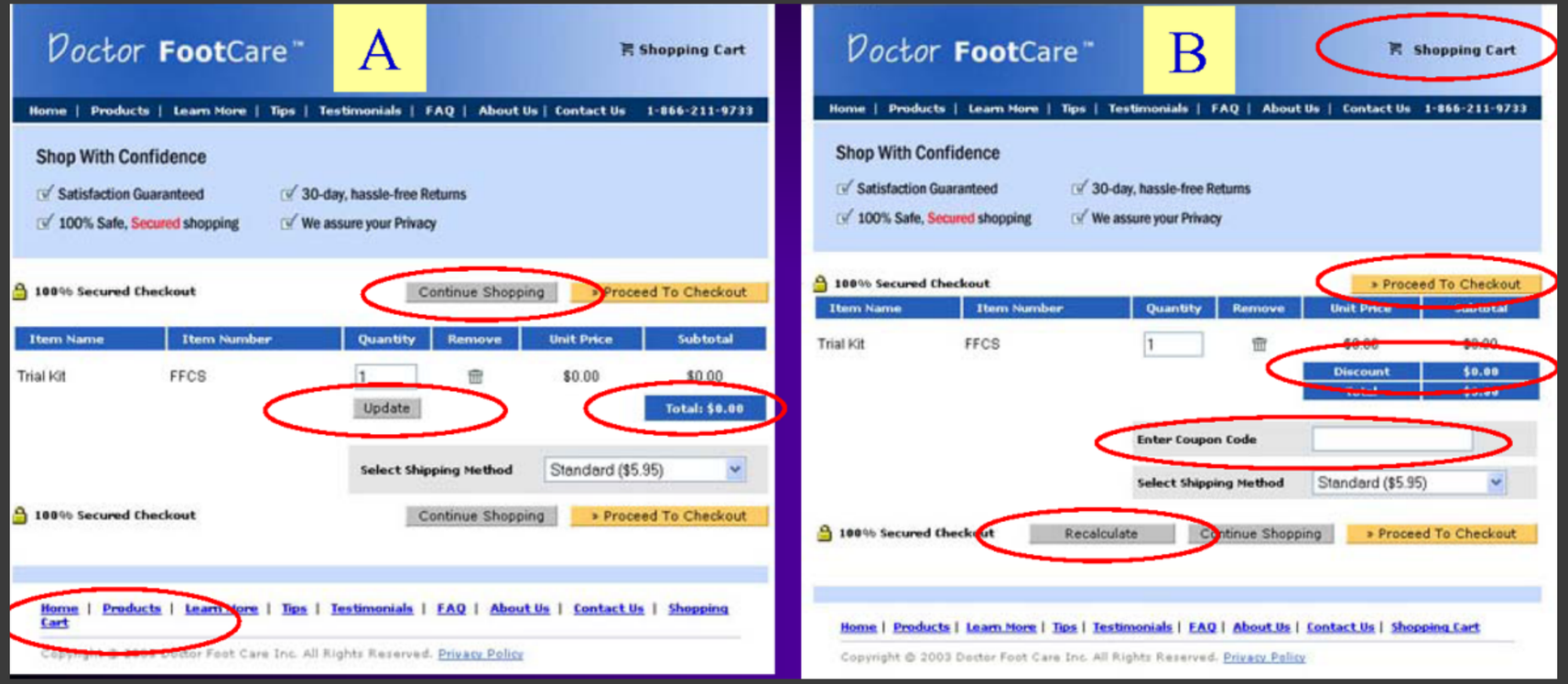
Content courtesy Ron Kohavi

# Large scale changes design

- Making small but consequential differences detectable.
- Small differences accumulate
- Beware of anomalies: investigate further

# Unexpected changes in a checkout page

- Conversion rate: the percent of visits to a Web site that include a purchase
- Which version has a higher rate?




courtesy Greg Linden's blog: http://glinden.blogspot.com/2006/04/early-amazon-shopping-cart.html

# The cost of one extra data field

By Nick Heath, 1 November 2010 16:15

**NEWS** Online travel firm Expedia has found that data analytics can deliver a multi-million dollar kick to a company's bottom line.

The company used SAS analytics to identify a single change to a web page that generated an overnight surge in sales, Expedia's VP of global analytics and optimisation Joe Megibow told the SAS Premier Business Leadership Series conference in Las Vegas last week.

Expedia analysts realised the site needed to be changed after investigating why many customers who clicked the 'Buy Now' button on the company's site did not complete the transaction.

"This is someone who was on our site, found the right location and hotel, put in all their billing and travel information and clicked the 'Buy Now' button," Megibow said.

"As far as leading indicators of purchase intent go, this is as good as it gets and yet we weren't taking the money."

Analysts began examining and correlating information about these failed transactions to identify what traits they had in common.

The answer, it turns out, was quite simple: "We had an optional field on the site under 'Name', which was 'Company'," Megibow said.

"It confused some customers who filled out the 'Company' field with their bank name."

FACEBOOK

silicon.com is on Facebook

facebook

For all the latest stories and updates from Silicon Towers...
'Like' us on Facebook

just film scott with the website in th background

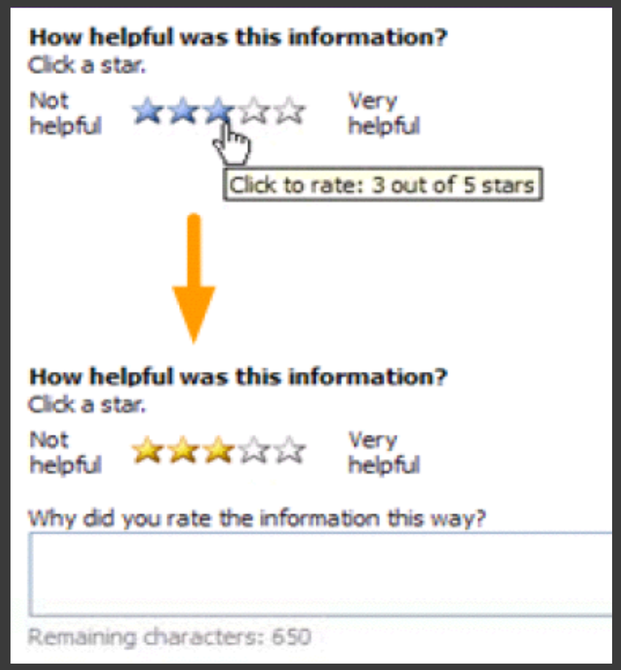
www.silicon.com/management/sales-and-marketing/2010/11/01/expedia-on-how-one-extra-data-field-can-cost-12m-39746554

Small distractions
*like extra fields*
can yield big changes

# Office Online Feedback

A

B




- Feedback A puts everything together, whereas feedback B is two-stage: question follows rating
- Feedback A just has 5 stars, whereas B annotates the stars with “Not helpful” to “Very helpful” and makes them lighter

**Which version has a higher response rate? By how much?**



courtesy Greg Linden’s blog: http://glinden.blogspot.com/2006/04/early-amazon-shopping-cart.html

**Please let us know if this content was helpful.**

Rate this content:

☆☆☆☆☆

Tell us why you rated the content this way (optional):

Remaining characters: 650

Submit

**How helpful was this information?**
Click a star.

Not helpful ★★★☆☆ Very helpful

Click to rate: 3 out of 5 stars

**How helpful was this information?**
Click a star.

Not helpful ★★★☆☆ Very helpful

Why did you rate the information this way?

Remaining characters: 650

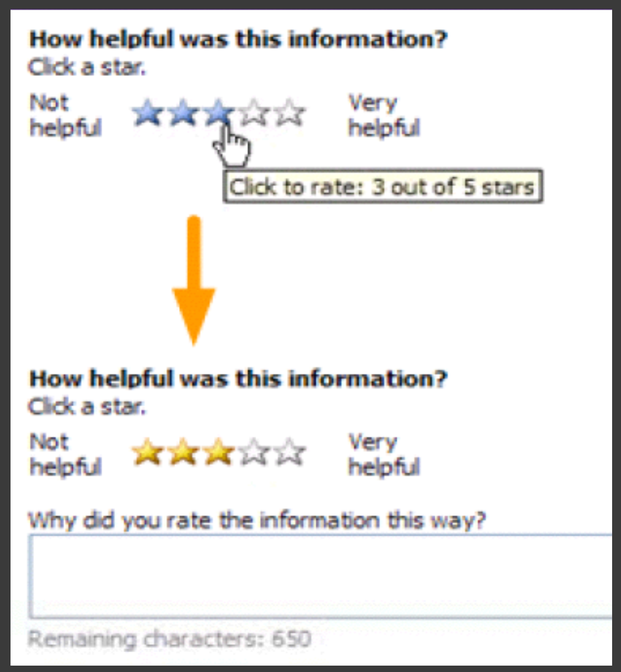
Submit

# Office Online Feedback

A

everything together just has five stars

B




two-stage: question follows rating annotates stars w/ “Not helpful” to “Very helpful”; makes them lighter

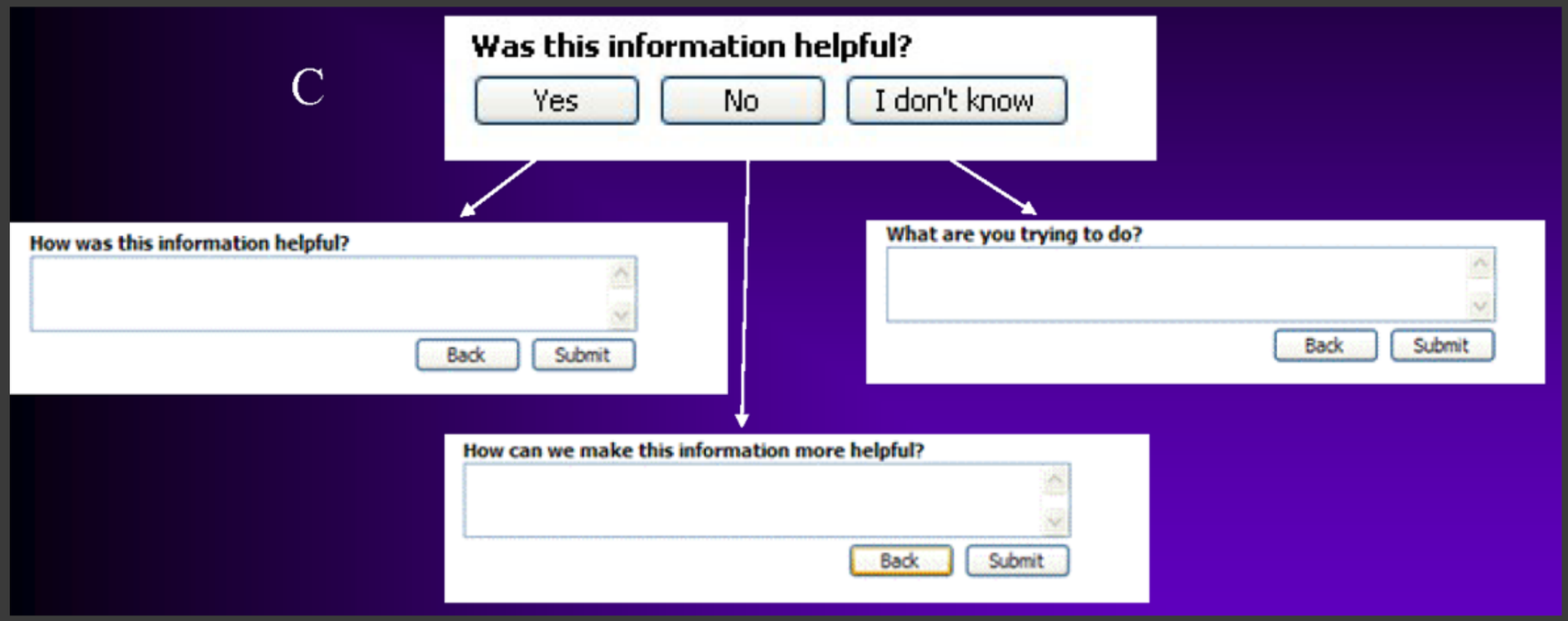
Which version has a higher response rate? By how much?

**B gets more than double the response rate!**



courtesy Greg Linden’s blog: http://glinden.blogspot.com/2006/04/early-amazon-shopping-cart.html

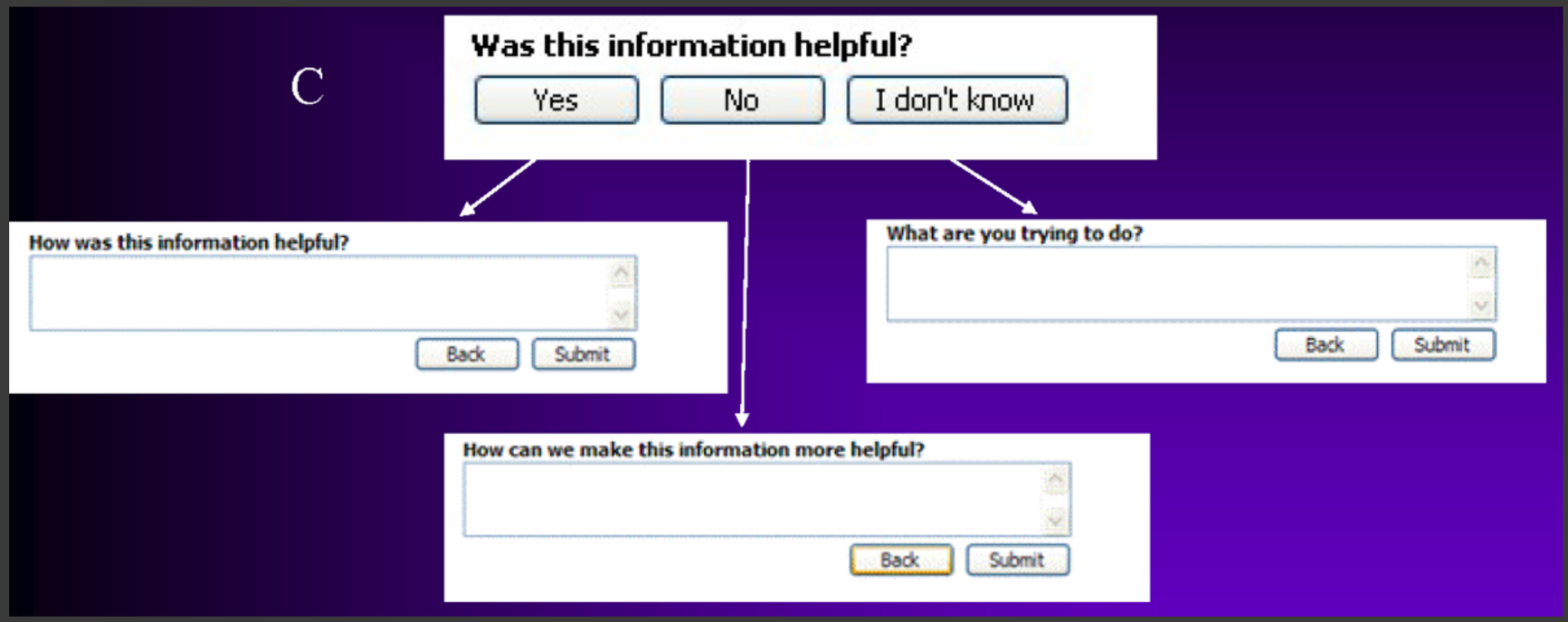
# Fewer options; custom response




courtesy Greg Linden's blog: http://glinden.blogspot.com/2006/04/early-amazon-shopping-cart.html

# Another Feedback Variant

C




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

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

**C outperforms B by a factor of 3.5!!**



courtesy Greg Linden's blog: http://glinden.blogspot.com/2006/04/early-amazon-shopping-cart.html

# Commitment Escalation

- Iterative design plus controlled experiments can enable you to fine-tune this stuff

# Run Experiments at 50/50%

- Novice experimenters run 1% experiments
- To detect an effect, you need to expose a certain number of users to the treatment (based on power calculations)
- Fastest way to achieve that exposure is to run equal-probability variants (e.g., 50/50% for A/B
- But don’t start an experiment at 50/50% from the beginning: that’s too much risk

  Ramp up over a short period



Content courtesy of : Ronny Kohavi

# Ramp-up and Auto-Abort

- **Ramp-up**
  - start an experiment at 0.1%
  - do some simple analyses to make sure [illegible] can be detected
  - ramp up to a larger percentage, and re[illegible]

Fix layout. Also, need right link for Kohavi.

- **Big differences are easy to detect beca[illegible] is quadratic in the effect we want to detect**
  - detecting 10% difference requires a small sample and serious problems can be detected during ramp-up
  - detecting 0.1% requires a population 100^2 = 10,000 times bigger

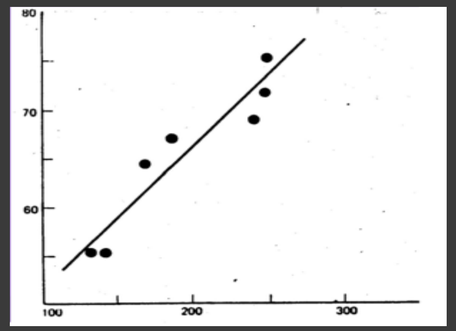
- **Automatically abort the experiment if treatment is significantly worse on OEC or other key metrics (e.g., time to generate page)**

Courtesy Ron Kohavi, Stanford HCI Seminar http://hci.stanford.edu/courses/cs547/speaker.php?date=2009-05-08

# Pick the a meaningful yardstick

# Doing the experiment is important because Correlations are not Necessarily Causal

- City of Oldenburg, Germany
- X-axis: stork population
- Y-axis: human population




What your mother told you about babies when you were three is still not right, despite the strong correlational “evidence”

- Example 2:

True statement (but not well-known):

Palm size correlates with your life expectancy

The larger your palm, the less you will live, on average

Try it out - look at your neighbors and you’ll see who is expected to live longer.

Why?

Women have smaller palms and live 6 years longer on average



Content courtesy of : Ronny Kohavi

# Run it for long enough

- Sometimes, first use is not the same as what people are familiar with

# Rules for Random Assignment

- Consistent
- Durable
- Independent

# Summary: Challenges & Opportunities

- Larger theories are still hard
- Using multiple methods together helps

# Design in the online age

- Designers role shifts to being about creating multiple alternatives
- People are often too sure of themselves
- Rapid experimentation means the first release is (sometimes) less important -- fail fast

# Turn to your neighbor

- When you run a Web experiment in A9
  - What might you measure?
  - What might you hope to learn?

# Announcements

- Extra Credit
- A7 has an example
- Decaf