

or other geometric designs



pedestal up
block on object

Alan's idea:
attach the wires to
the middle tubing

Web-Scale Experimentation

Scott Klemmer
Autumn 2009



Scott: a gate that shows
who walked through it last
Bill: a gate that measures
ceremonial gates



passive gate
array



pinball
machine

Controlled Experiments on the Web

- Many names; same idea
 - A/B tests
 - Randomized experiment
 - Controlled experiment
 - Split testing
- Randomly split traffic between two versions
 - A/Control: usually current live version
 - B/Treatment: new idea
- Collect metrics, analyze



GET INVOLVED



← Media

JOIN THE
MOVEMENT

Email Address:

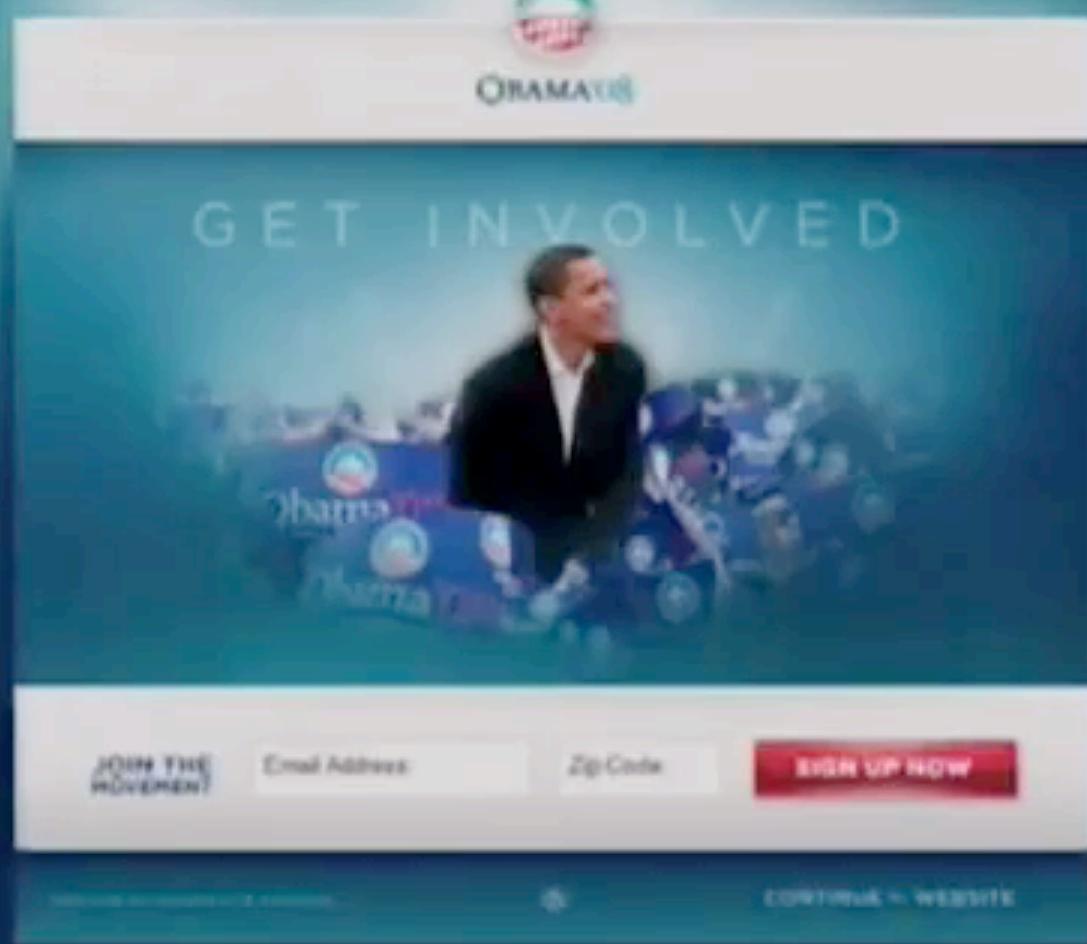
Zip Code:

SIGN UP

← Button

Baseline sign up rate: 8.26%

STANFORD
UNIVERSITY



Button: "Sign Up Now"

STANFORD
UNIVERSITY





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:	<input type="text"/>
(Email address used for RED ALERT Notifications)	
First Name:	<input type="text"/>
Last Name:	<input type="text"/>
Address 1:	<input type="text"/>
Address 2:	<input type="text"/>
City:	<input type="text"/>
State:	<input type="select" value="Select"/>
Zip:	<input type="text" value="32225"/>
Country:	<input type="select" value="United States"/>
Phone:	<input type="text"/>

Billing Information

By submitting this information you agree to a one time charge of \$10.00. [Terms and Conditions.](#)

Upgrade my registration so I can generate "Unlimited Reports" when new predators are in my area for just \$4.95/month. Yes No



Credit Card Number:

Credit Card Exp. Date: /

3 or 4 Digit Code After Card # on Back of Card:

[CLICK HERE FOR INSTANT ACCESS](#)

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. 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 offenders**, including specific addresses. The full service search is based on **your exact address**.

From <http://www.anspaper.com/articles/designcaricriptie>

However compelling the message, however great the copy, however strong the sales argument... the way a page is designed will have a dramatic impact on conversion rates, for better or for worse. Here are three versions of the same offer page I know, they won't win any design awards. They weren't intended to. But they are functional and familiar. A reader going to any one of these pages will be able to quickly figure out what the message is, and what they are being asked to do.

Version A is the original.

Search Results for Zip Code 32225

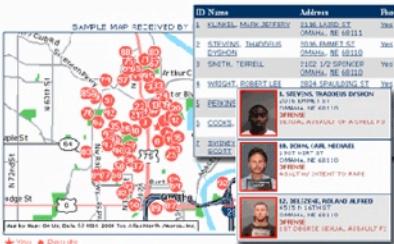
Search Initiated:	8/2/2005	Time:	12:11 P.M.
Search Location:	State: FL	City: JACKSONVILLE, JAX	Zip: 32225
There are currently 32 Registered Sex Offenders located in your zip code			

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 custom 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**



You 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: <input type="text"/>
Last Name: <input type="text"/>
Address 1: <input type="text"/>
Address 2: <input type="text"/>
City: <input type="text"/>
State: <input type="select" value="Select"/>
Zip: <input type="text" value="32225"/>
Country: <input type="select" value="United States"/>
Phone: <input type="text"/>

Billing Information

By submitting this information you agree to a one time charge of \$10.00. [Terms and Conditions.](#)

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

Upgrade my registration and guarantee me "Red Alert" **Automatic Report Updates & Email Notifications** when new predators are in my area for an additional \$4.95/month. [Details](#)



Credit Card Number: 

Credit Card Exp. Date: /

3 or 4 Digit Code After:

Card # on Back of Card: [How to find your security code?](#)

[CLICK HERE FOR INSTANT ACCESS](#)



"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



From <http://www.alistapart.com/articles/designcancripple>

Version B follows the same basic layout, but we made some minor copy changes.

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
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	<input type="button" value="▼"/>
Zip:	32225	
Country:	United States	<input type="button" value="▼"/>
Phone:		

[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](#)

Optional **Red Alert** Program 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.

"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.**"
Tolene 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."

"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*

From <http://www.alistapart.com/articles/designcancripple>

In version C, we changed from a one-column format to two-column format. We wanted to test the impact of bringing more of the page content onto the first screen.

Be honest with yourself and decide now whether B or C beat A, and by what percentage. I imagine you have some way of measuring the success of your site. Maybe it's about sales. Maybe it's based on readership. But one way or another, your site has a purpose.

But I don't think most designers truly understand the effect their design choices can have on achieving that purpose.

And yes, I'm sure you do some usability testing. And that likely gives you some broad, if sometimes confusing insights into what's working and what isn't.

But do you test different page designs?

By testing, I don't mean asking a few folks around the office; I mean doing a live test that demonstrates—with hard figures—what site visitors actually do.

Testing like that is a beautiful thing. There is no space for fancy arguments. An expert's credentials and opinions mean squat. When you serve alternative versions, one after the other, and measure reader actions, you get the real deal. You get what is. But if you are serious about achieving your site's purpose, and if testing can show you which version of a page does best, then where is the argument not to test?

Ways design makes a difference

- The position and color of the primary call to action
- Position on the page of testimonials, if used
- Whether linked elements are in text or as images
- The amount of “white space” on a page, giving the content space to “breathe”
- The position and prominence of the main heading
- The number of columns used on the page
- The number of visual elements competing for attention
- The age, sex and appearance of someone in a photo

From <http://www.alistapart.com/articles/designcancripple>

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%

From <http://www.alistapart.com/articles/designcancripple>

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](#)



[Don't Kill Projects; Morph Them](#)

[Marissa Mayer](#)



[Taking Risks](#)

[Vinod Khosla](#)

2 min. 1 sec.

2 min. 2 sec.

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 ecosystem for all.



Podcast: [Healthy Entrepreneurship in Medical Devices](#)

[Mir Imran](#), InCube Labs - 54 min. 28 sec.

Parallel entrepreneur Mir Imran, CEO of InCube Labs, has launched twenty companies - at times simultaneously. He shares his solutions-focused expertise and identifies the vitality and growth of the biomedical healthcare vertical.



Podcast: [The Path to an Acquisition](#)

[Brett Crosby](#), Google Analytics - 1 h 1 min. 5 sec.

Brett Crosby, Group Manager of Google Analytics, describes the ebb and flow of the process by which his web analytics company, Urchin, was acquired by Google. He also shares some inspirational lessons in making small business loom large.

Educators Corner, a project of [Stanford Technology Ventures Program](#) at Stanford University, is a collection of recordings from hundreds of Silicon Valley's most practiced entrepreneurs and thought leaders. The collection offers over 1,200 videos and podcasts, freely available to the public.

Please [send us an email](#) if you have any comments.

Terman Engineering Center, 380 Panama Mall, Stanford, CA 94305

Click [here](#) to unsubscribe.

From Forrest Glick, Stanford STVP <http://project8180.stanford.edu/category/wireframes/>

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 unreasonably. Entrepreneurs need to stretch themselves beyond convention and constraint to reach something extraordinary.

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 ecosystem for all.



Podcast: [Healthy Entrepreneurship in Medical Devices](#)

[Mir Imran](#), InCube Labs - 54 min. 28 sec.

Parallel entrepreneur Mir Imran, CEO of InCube Labs, has launched twenty companies - at times simultaneously. He shares his solutions-focused expertise and identifies the vitality and growth of the biomedical healthcare vertical.

From Forrest Glick, Stanford STVP <http://project8180.stanford.edu/category/wireframes/>

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

From Forrest Glick, Stanford STVP <http://project8180.stanford.edu/category/wireframes/>

NY Times, July 1, 2009

Facebook to Offer New Features to Allow Users to Control Privacy of Information

By **JENNA WORTHAM**
Published: July 1, 2009

Facebook's biggest asset — its population of more than 200 million users — is also part of its Achilles' heel. As more people join and connect with more people they know, the chances grow that one's embarrassing photographs from a night of carousing might be seen by the wrong person.

Related

Times Topics: [Facebook](#)

To help remedy this, the company is testing new controls that will allow members to specify which groups or individuals are able to see each text update, photo or video they post on the site. For example, the controls would make it easy for a user to remind all the family members on Facebook about a surprise birthday party, while excluding the birthday girl.

"Our overall philosophy is that people should be as open or as closed as they want to be," said Chris Kelly, chief privacy officer for the company, in a conference call with reporters Wednesday to discuss the changes.

For now the new features are available only to a limited number of users, but the company said it planned to provide them eventually to everyone on the site.

The changes are part of Facebook's effort to simplify its privacy settings, which had ballooned to more than six pages and 40 different options. The company said that fewer than a quarter of its users regularly adjusted privacy settings, so it planned to condense them into a single, easy-to-navigate page.

- [SIGN IN TO RECOMMEND](#)
- [TWITTER](#)
- [E-MAIL](#)
- [SEND TO PHONE](#)
- [PRINT](#)
- [REPRINTS](#)
- [SHARE](#)

ARTICLE TOOLS
SPONSORED BY



Project goals (from Julie Zhuo)

- message the fact that privacy has been simplified, and there is now an ‘everyone’ option
- encourage people to open up the information they’re comfortable sharing in order to make themselves more discoverable in search
- make sure point 2 is done with full user understanding.

Facebook is making changes to privacy—update your settings now.

Facebook has simplified Privacy to make it easier to control your information. You can share some of your info with Everyone or choose to apply your old settings to the fields below. Note that Everyone means everyone on the internet.

	Everyone	Old Setting
Basic Info [?]	<input checked="" type="radio"/>	<input type="radio"/>
Family and Relationships [?]	<input checked="" type="radio"/>	<input type="radio"/>
Education and Work	<input type="radio"/>	<input checked="" type="radio"/>
Personal Info [?]	<input type="radio"/>	<input checked="" type="radio"/>
Photos and Videos of Me [?]	<input type="radio"/>	<input checked="" type="radio"/>
Wall tab and Posts I Publish [?]	<input type="radio"/>	<input checked="" type="radio"/>
Photos Tab	<input type="radio"/>	<input checked="" type="radio"/>

Q: Chris Berry

Displaying 1 – 10 out of over 500 results for: chris berry

	Name: Chris Berry	Network: New York, NY	Friends: 4 mutual friends
	Name: Chris Berry	Friends: 1 mutual friend	
	Name: Chris Berry	Friends: 1 mutual friend	

Friends can connect with you more easily when more information is available.

[Save Settings](#)

[Customize Settings](#)

Requests

[See All](#)

50 friend requests 1 event invitation

1 other request 1 new update

Suggestions

[See All](#)



Ben Newman

47 people are mutual friends. You both went to Stanford. [Add as Friend](#)

Sponsored

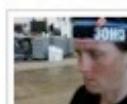
Only in a Woman's World



Join the four, funny ladies of A Woman's World for hilarious webisodes, great conversation and a chance to win a girls getaway to NYC.

[Become a Fan of Only in a Woman's World](#)

Highlights



Mobile Uploads
by Alison Rosenthal

1



Wall Photos
by Jared Morgenstern

3 11

[Disappearing Cat Door](#)

News Feed

campfire

designers

girls

platform

si ren



What's on your mind?

Attach

[Share](#)



Facebook Design [Rob] — Found the logo history between Pepsi and Coke thought-provoking. Coca-Cola's mass-culture brand is a testament to the power of consistency. Stick to your style — it may fall out of fashion from time to time, but if it's really good it will always come back.

Facebook is making changes to privacy—update your settings now.

Facebook has simplified Privacy to make it easier to control your information. Please choose your new privacy level from the options below. Note that Everyone means everyone on the internet:

Open

Everyone can see your profile and posts. [Show details](#) ▾

Recommended

Everyone can see your general info and posts. Friends of Friends can see your sensitive info. [Show details](#) ▾

Limited

Only friends can see your entire profile and content that you publish. [Show details](#) ▾

Apply Old Settings

Only people who could see your info before will continue to have access. No information or content will be exposed.

[Save Settings](#)

[Customize Settings](#)

Q: Chris Berry

Displaying 1 – 10 out of over 500 results for: chris berry

	Name: Chris Berry	Network: New York, NY	Friends: 4 mutual friends
	Name: Chris Berry	Friends: 1 mutual friend	
	Name: Chris Berry	Friends: 1 mutual friend	

Friends can connect with you more easily when more information is available.

Requests

[See All](#)

50 friend requests

31 1 event invitation

1 other request

1 new update

Suggestions

[See All](#)



Ben Newman

47 people are mutual friends.

You both went to Stanford.

[Add as Friend](#)

Sponsored



How strong is your score?



See your 2009 credit score online in just two easy steps from Experian.

Highlights



Audio slide show: GIVE HIM HIS PROPS – Los An...

posted by Alexandre Roche

2



Shooting Jan's Skate Session

Alexandre Roche likes this.



Oooold School Pics

Sasha Rush is tagged.

[News Feed](#)

campfire

designers

girls

platform



What's on your mind?

Attach



[Share](#)

Eston Bond Late. Very.

6 minutes ago · Comment · Like

News Feed

campfire

designers

girls

platform

si ren

Pages

Links

Video

Notes

Photos

Typing Maniac

Status Updates

FriendFeed

More



What's on your mind?

Important Privacy Announcement

Facebook is making changes to privacy—update your settings now.

Facebook has simplified Privacy to make it easier to control your information. You can share some of your info with Everyone or choose to apply your old settings to the fields below. Note that Everyone means everyone on the internet.

Everyone Old Setting

Basic Info [?]



Family and Relationships [?]



Education and Work



Personal Info [?]



Photos and Videos of Me [?]



Wall tab and Posts I Publish [?]



Photos Tab



Q+ Chris Berry

Displaying 1 – 10 out of over 500 results for chris berry

	Name: Chris Berry	Network: New York, NY	Friends: 4 mutual friends
	Name: Chris Berry	Friends: 1 mutual friend	
	Name: Chris Berry	Friends: 1 mutual friend	

Friends can connect with you more easily when more information is available.

Save Settings

Skip For Now

Customize Settings



Babak Irani Old one from NikAhang Kosar:
Ayatollah 'Temsah' (Crocodile) (pun on Ayatollah Mesbah): Will anyone rescue me from this sell out journalists?

www.muslim-refusenik.comSource: www.muslim-refusenik.com

by Jared Morgenstern

3 11



Disappearing Car Door posted by Nirav Sanghani

3 2



News Feed

campfire

designers

girls

platform

si ren

Pages

Links

Video

Notes

Photos

Typing Maniac

Status Updates

Microsoft

More



What's on your mind?

Important Privacy Announcement

Facebook is making changes to privacy—update your settings now.

Facebook has simplified Privacy to make it easier to control your information. Please choose your new privacy level from the options below. Note that Everyone means everyone on the internet:

 Open

Everyone can see your profile and posts. [Show details ▾](#)

 Recommended

Everyone can see your general info and posts. Friends of Friends can see your sensitive info. [Show details ▾](#)

 Limited

Only friends can see your entire profile and content that you publish. [Show details ▾](#)

 Apply Old Settings

Only people who could see your info before will continue to have access. No information or content will be exposed.

Requests

50 friend requests

1 event invitation

1 new update

See All

Q: Chris Berry

Displaying 1 – 10 out of over 500 results for: chris berry



Name: Chris Berry
Network: New York, NY
4 mutual friends



Name: Chris Berry
Friends: 1 mutual friend



Name: Chris Berry
Friends: 1 mutual friend

Friends can connect with you more easily when more information is available.

Save Settings

Skip For Now

Customize Settings



Eston Bond Late. Very.

7 minutes ago · Comment · Like

Share · Comment · Like



Mega Rally SF, July 25th
(Part 2)

posted by Babak Irani



News Feed

campfire

designers

girls

platform

si ren

Pages

Links

Video

Notes

Photos

Typing Maniac

Status Updates

Microsoft

More



What's on your mind?

Important Privacy Announcement

Facebook is making changes to privacy—update your settings now.

Facebook has simplified Privacy to make it easier to control your information. Please choose your new privacy level from the options below. Note that Everyone means everyone on the internet:

 Open

Everyone can see your profile and posts. [Show details ▾](#)

 Recommended

Everyone can see your general info and posts. Friends of Friends can see your sensitive info. [Show details ▾](#)

 Limited

Only friends can see your entire profile and content that you publish. [Show details ▾](#)

 Apply Old Settings

Only people who could see your info before will continue to have access. No information or content will be exposed.

Requests

50 friend requests

1 event invitation

1 new update

See All

Q: Chris Berry

Displaying 1 - 10 out of over 500 results for: chris berry

	Name: Chris Berry
Network: New York, NY	Friends: 4 mutual friends
	Name: Chris Berry
Friends: 1 mutual friend	
	Name: Chris Berry
Friends: 1 mutual friend	

Friends can connect with you more easily when more information is available.

Save Settings

Skip For Now

Customize Settings



posted by Babak Irani

July 25th



BEAR ATTACK

by Evan Priestley

+50

Help your friends find you on Facebook.

We've simplified the privacy settings. Now you can make it easier for friends to connect with you by making a few fields visible to everyone on the internet. Please take a moment to review our recommended settings:

Profile Info, Wall and Status Updates

 Everyone ▾

Photos and Videos

 My Friends and Their Friends ▾

Private Info and Contact Info

 Only My Friends ▾

Your custom settings will be preserved for:

- Videos tagged of you, Interests, and Email addresses

Save Settings

or for more information, view your [Privacy Settings](#)

So the best option out of the 5 appears to be modal popup with buckets and no preselected states.

Lesson: Compute Statistical Significance and run A/A Tests

39

- A **very common mistake** is to declare a winner when the difference could be due to random variations
- **Always run A/A tests**
(similar to an A/B test, but besides splitting the population, there is no difference)
- **Compute 95% confidence intervals on the metrics to determine if the difference is due to chance or whether it is statistically significant**
- **Increase percentage** if you do multiple tests (e.g., use 99%)
- **Idea:** run an A/A test concurrently with your A/B test to make sure the overall system doesn't declare it as significant more than 5% of the time (great QA)



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

Ramp-up and Auto-Abort



- **Ramp-up**
 - Start an experiment at 0.1%
 - Do some simple analyses to make sure no egregious problems can be detected
 - Ramp-up to a larger percentage, and repeat until 50%
- **Big differences are easy to detect because the min sample size 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)**

Randomization

- **Good randomization is critical.**

It's unbelievable what mistakes devs will make in favor of efficiency



- **Properties of user assignment**

- Consistent assignment. User should see the same variant on successive visits
- Independent assignment. Assignment to one experiment should have no effect on assignment to others (e.g., Eric Peterson's code in his book gets this wrong)
- Monotonic ramp-up. As experiments are ramped-up to larger percentages, users who were exposed to treatments must stay in those treatments (population from control shifts)

Two-sample t -test

- *Signal/Noise*
- X = observed mean (test and control)
- var = variance
- n = sample size

$$\frac{\bar{X}_T - \bar{X}_c}{\sqrt{\frac{\text{var}_t}{n_t} + \frac{\text{var}_c}{n_c}}}$$

Example con't

$$\frac{12s - 10s}{\sqrt{\frac{2.0s}{16}} + \sqrt{\frac{2.0s}{16}}} df = 16 + 16 - 2 = 30$$

Simplify...

$$t = 4 df = 30$$

Example, con't

- Use SPSS/R/SAS, or use a table:
- Use your df , the p value you're aiming for, and the t value you found

