“Do You Like My Interface?”
“How much do you like my interface?”

“This is a useful interface: agree/disagree”
Please the Experimenter Bias
However:
Watching people fail to use your design is one of the most effective learning tools anywhere.
Today: measuring success offline and online

- Variables, validity
- Randomization
- Usability testing
- Online A/B testing
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
A Not-Very-Useful iPhone Keyboard Study

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.
Benefits and Drawbacks?

- Manipulation: Input Style
- Measure: Words per minute
- External validity: not so much
A better version: actual users

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

Figure 1. Average time to type a message on phones owned by the participants (M +/- SD).
iPhone & Qwerty users similar speed, but

Figure 1. Average time to type a message on phones owned by the participants ($M \pm SD$).

Figure 2. Average number of total errors per message made by participants using their own phones ($M \pm SD$).

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
1) Experimental Design
Controlled comparison enables causal inference.

Between vs. within subjects designs
Randomization
Should every participant use every alternative?
Which professor style is more effective?
What are the measures?

· Faster?
· Better test scores?
· Fatigue?
· Attention focused on lecture?
· End-of-class evaluations?
• Manipulation: Professor style
• Measure: Test scores, end-of-class evaluations
Between subjects design

Half the participants see one version

The other half see the other
Within subjects design

Everyone uses both versions
How Can We Address Ordering Effects?
How about individual differences?

Randomization washes them out.
What about for Three or More Alternatives?
# Latin Square

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
What happens if we don’t randomize?
Self-selection effects

- Typing in the morning versus the afternoon
Learning effects

- Showing alternatives in sequence
At a high level:
Should every participant use every alternative?
Two Major Strategies

· **Within-subjects:** everyone tries all the options. Good when you’re not worried about learning/practice/exposure issues (that trying one version will ‘pollute’ the date from another version)

· **Between-participants:** each person tries one. Requires more people, and more attention to fair assignment. Has the benefit that each participant is uncorrupted (at least by the study...)

37
2) In-person experiments
Set a clear, focused goal

• Scope: make a meeting room booking system for Gates.
• Purpose: create a system that encourages people to not overbook the length of time they need.
• Hypothesis: splitting the booking process over several screens will encourage more thought and people will book better.
Recruit to match that goal

- Schedule and location: next week, Stanford Gates building
- Participants: 12 people (4 students, 4 office administrators, 4 professors)
Create concrete tasks

“Book a room sometime next week for a research group meeting. Andrew will be out of town, so we won’t hear his weekly update. The rest of us should be present and give our updates. Besides the usual group members, we’ll have two visitors from France who will present their research—maybe they’ll take 10 minutes each.

When you’re done booking the room, tell Arvind so he can prepare the next task for you.”
“We are testing our design: we are not testing you.”
Experimental Details

- Task ordering
  - Start simple, then build up to complex tasks
- Training: is this a walk-up-and-use system? Or will real users receive training?
- What if someone doesn’t finish?
Always pilot your study before launching

- A pilot study is a practice run before the experiment
  - Pilots = low-cost prototypes of experimental design
  - Even if you are running behind schedule, do a pilot study
- Why pilot?
  - Debug study protocols
  - Catch errors early so they don’t mess up your results
- Run two pilots
  - First: friends+colleagues
  - Second: real users
Options for capturing results

- Bring a notebook for freeform feedback
- Instrumented software
- Video recording
- Screen recording
Think Aloud protocol

• Need to know what users are thinking, not just what they are doing
• Ask users to talk while performing tasks
  • tell us what they are thinking
  • tell us what they are trying to do
  • tell us questions that arise as they work
  • tell us things they read
• Prompt the user to keep talking:
  • “Tell me what you are thinking”
• Make a recording or take good notes
  • make sure you can tell what they were doing
3) Online field experiments
Relative merits: online vs. in-person

Offline: deep understanding
Online: scale and resolution
National Alert Registry

Search Results

Search Initiated: 6/30/2004  Time: 12: 12 AM
Search Location: Zip: 32005 City: JACKSONVILLE, FL

32 Registered sex offenders are located in your immediate area.

Set the complete report, including map, photograph, and address 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:

- Address
- Appearance Details
- Criminal Information
- Offense Dates
- Offense Details
- Street Address
- Name

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 "No Risk" notification when new predators are in your area.

Last Name:
Address 1:
Address 2:
City:
State:
Zip:

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 "Individual Reports" when new predators are in my area for just $10.00.

Yes

Credit Card Information:

Credit Card Type:
2 or 4 Digit Code:

Submit

Interface courtesy of A List Apart: http://www.alistapart.com/articles/designcancripple
Interface courtesy of A List Apart: http://www.alistapart.com/articles/designcancripple
Create Your Sex Offender Map And Detailed Predators 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 $1.50.

Your Neighborhood Report

We will use your address to create a custom report:

- Email:
- First Name:
- Last Name:
- Address 1:
- Address 2:
- City:
- State:
- Zip Code:
- Country:
- United States:

Select your billing preference:

- Single "Predator Report" for a one-time charge of $1.50.
- Upgrade my registration to "Red Alert." Only $9.95/month now.

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 nationwide database has been updated.

Credit Card Information:

- Credit Card Type:
- Credit Card Number:
- Credit Card Exp. Date:
- Security Code:

SUBMIT NOW FOR INSTANT ACCESS

The complete Predator Report includes:

- Color Photo(s) (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 nationwide database has been updated.

"Your Predator Report was created by Safe Zone, a national group committed to helping people in your area protect themselves and their families from child predators. Safe Zone has a nationwide network of thousands of volunteers who operate in your area.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.

We are not affiliated with Safe Zone as it is a business name for a different company.
## A/B/C Split Test

<table>
<thead>
<tr>
<th></th>
<th>PAGE A</th>
<th>PAGE B</th>
<th>PAGE C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of traffic</td>
<td>34%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>New sales</td>
<td>244</td>
<td>282</td>
<td>114</td>
</tr>
<tr>
<td>Change</td>
<td>N/A</td>
<td>15.57%</td>
<td>-53.28%</td>
</tr>
</tbody>
</table>
Free Search Results for Zip Code 32225

Search Initiated: 9/2/2005  
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

Sample Report

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 whenever 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.

Optional Red Alert Program includes:

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

See the entire sample Report!
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
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

Vined 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 Houlihan, B Corporation and Andrew Kassey, B Corporation – 50 min. 55 sec.

The co-founders of B Lab, Jay Coen Gilbert, Bart Houlihan, and Andrew Kassey, 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.
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 jet to your fast-tracking career.

Video: Negotiating with Customers and Clients
Stan Ovshinsky, 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 anew set of wings. With Google’s 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 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 Duggan, 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.

Courtesy 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
Which option performed best?

- Sign up
- Learn More
- Sign up Now
- Join Us Now
Now for the visual material: five options
Button: “Sign Up Now”
Which do you think did best?

• Family Image
• Change Image
• Barack Video
• Springfield Video
• Sam’s Video
Here We Saw

• Small changes: big difference
• Our expectations are often wrong
Which got the most clickthroughs?

• I’m on twitter
• Follow me on twitter
• You should follow me on twitter
• You should follow me on twitter here
Which got the most clickthroughs?

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
Typography Experiment:
Color Contrast on MSN Live Search

A: Softer colors

B: High contrast
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
- Watch out for spurious results: do you trust your measurements?
Unexpected changes in a checkout page

- Which version produces more purchases?
The cost of one extra data field

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.
Small distractions such as extra fields can yield big changes
Small changes have positive impact, too

Version A

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

Version B

2x the response rate of Version A

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

Version C

3.5x the response rate of Version B!
Why?

• Commitment escalation: if they agree to do a little bit, then add more later, they’re much more likely to do it than if you ask for everything up front
The big picture: continuous iteration

- Iterative design plus controlled experimentation are a formidable combination for quickly improving your design
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