IN4MATX 231:
User Interface Design & Evaluation

Class 12:
High-Fidelity Prototyping & Summative Evaluations

Daniel Epstein

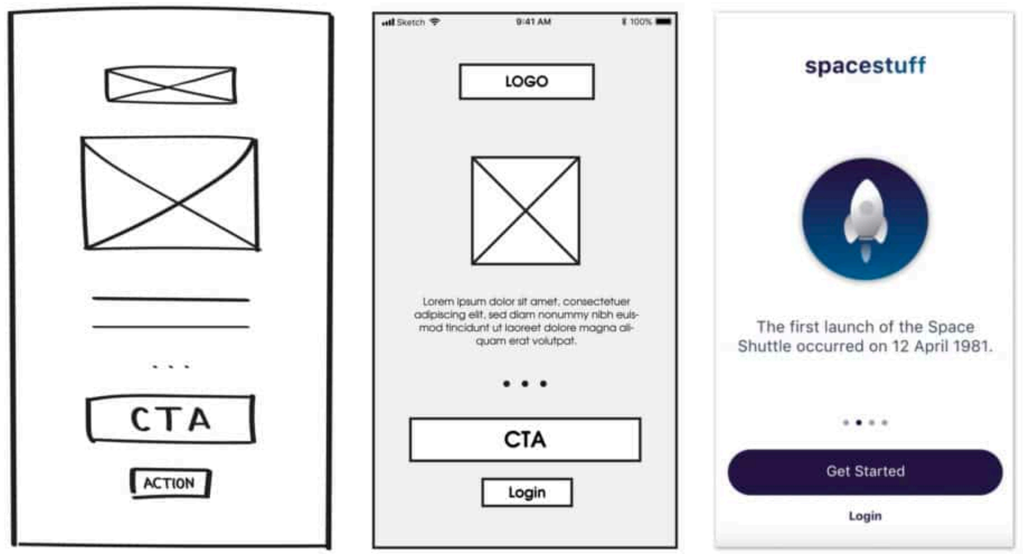
Today’s goals

By the end of today, you should be able to…

- Develop useful prototypes at higher fidelity than paper
- Think through decisions about study size and environment in summative evaluation
- Make choices about study length, style, and population in a field evaluation

Fidelity in Prototyping

- High Fidelity
  - Prototypes which look like the final product
- Low Fidelity
  - Designer sketches with many details missing

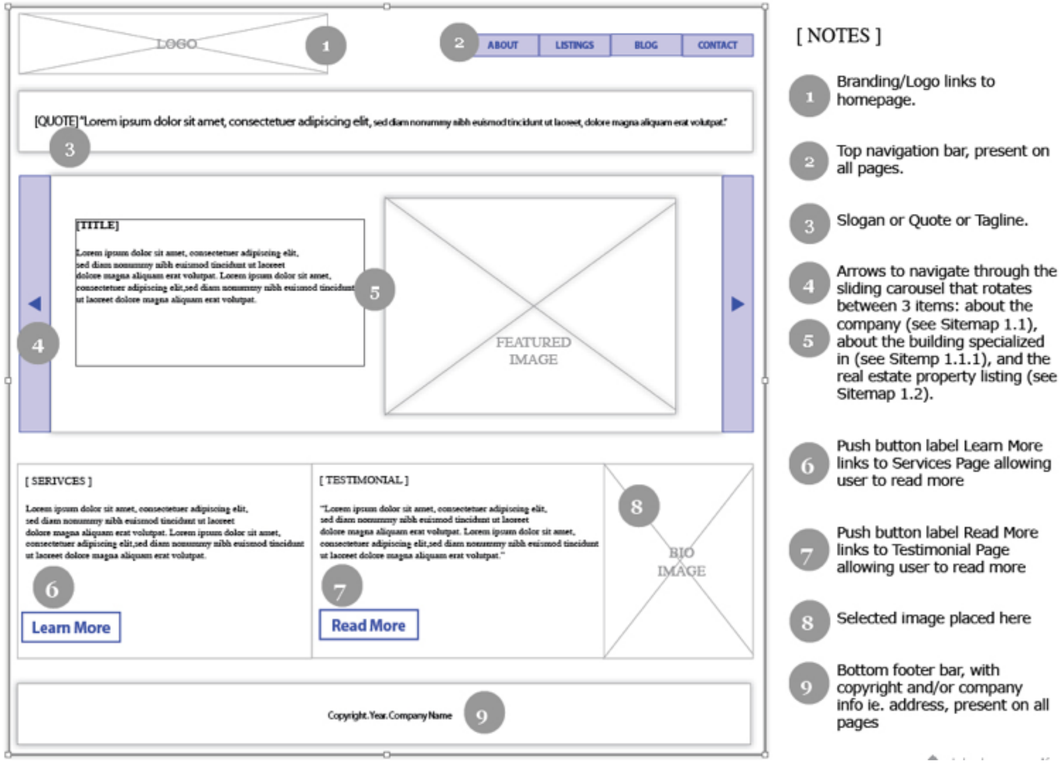


High Fidelity Prototyping

- Who is your audience?
  - Probably your design team
  - Maybe whoever is implementing your design
  - *Possibly* a potential user
- How will they be used?
  - In a design critique
  - As implementation guides

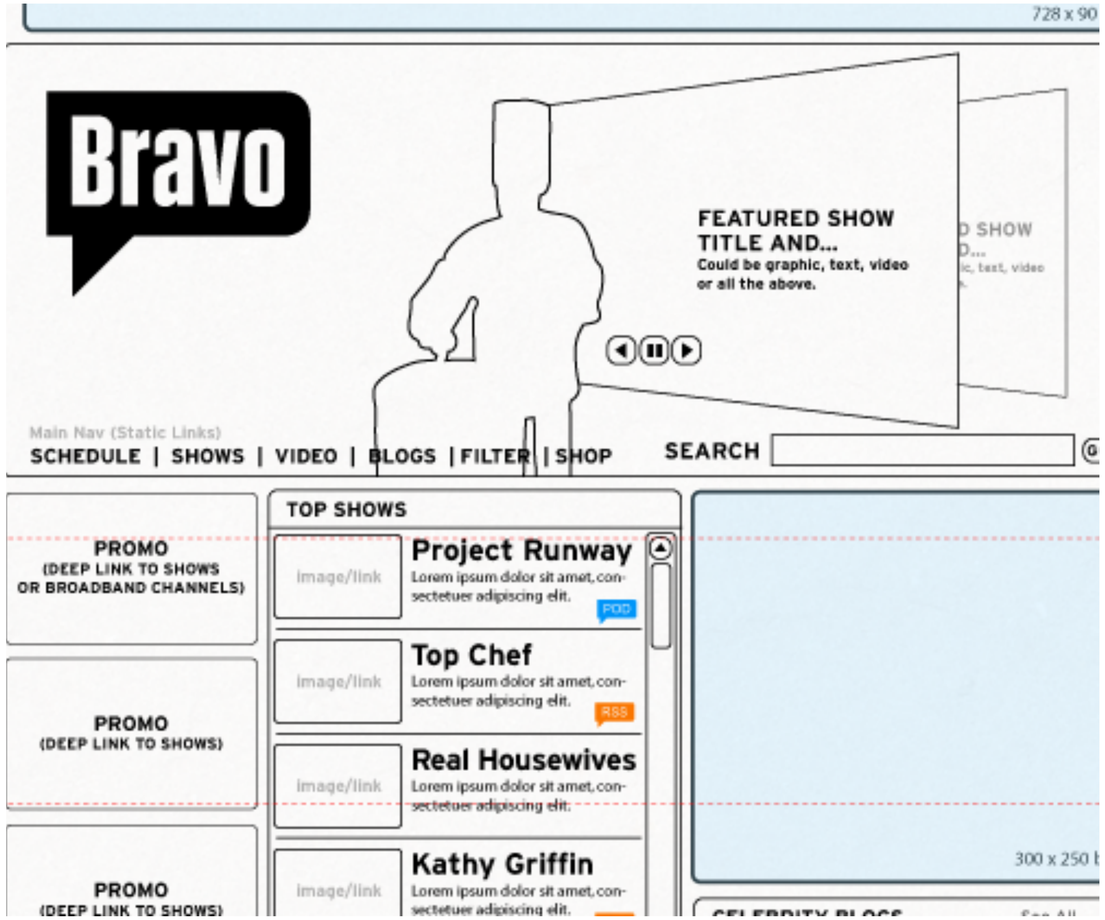
High Fidelity Prototyping

- Wireframes include layouts and key functional units
- They do not contain visual design details



High Fidelity Prototyping

- Wireframes include layouts and key functional units
- They do not contain visual design details

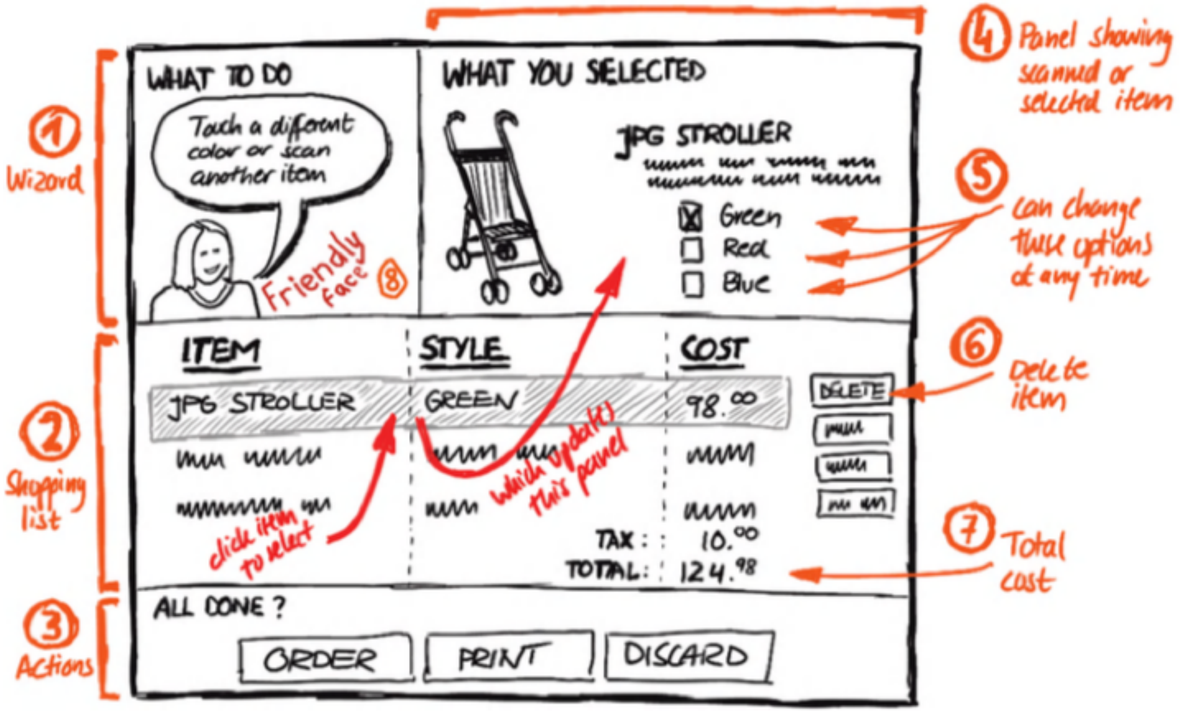


High Fidelity Prototyping

- Use greeking
  - Or text blocks
  - Or placeholders instead of images

High Fidelity Prototyping

- Annotate your wireframes
  - Good annotations clearly indicate actionable items



Annotation Best Practices

- Keep them short and to the point
  - “Automatically-generated photo from most recent blog post”
- Focus on user benefits
- Use numerical circles to order them (flags are good, too)
- Locate explanations to the right of the wireframe
  - People will scan for annotation circles on the left, move to the right to read explanation
- Use consistent terminology for objects

Annotation Examples

- *“Triggering this call to action shall result in the display of the home page”*
- *“What the Events tab will look like when it’s selected by the user. It will have an icon and show the number of events.”*
- *“Name of the Event in bold. Below is a short description of the event.”*
- *“‘Attending’ flag that appears when a user RSVPs attending.”*
- *“To RSVP for an event, users have to click this button. It shows how many days until an event.”*

Annotation Examples

- *“Will link to view the details of the random example displayed from the portfolio of work.”*
- *“Title of the most recent post.”*
- *“Will link to view the full blog post of the most recent live blog posting.”*
- *“Shall link to the Terms & Conditions page from any location within the website”*
- *“Image of the random example displayed from the portfolio of work. Shall link to details of the random example displayed from the portfolio of work.”*
- Formality supports credibility and authority

Annotating your vending machine

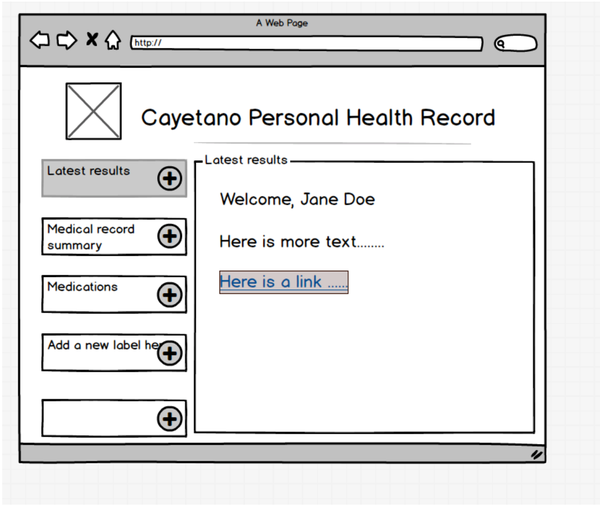
- Write annotations for the paper prototype you made for your vending machine
- We will discuss as a group what questions we might have if we were in charge of implementing your design

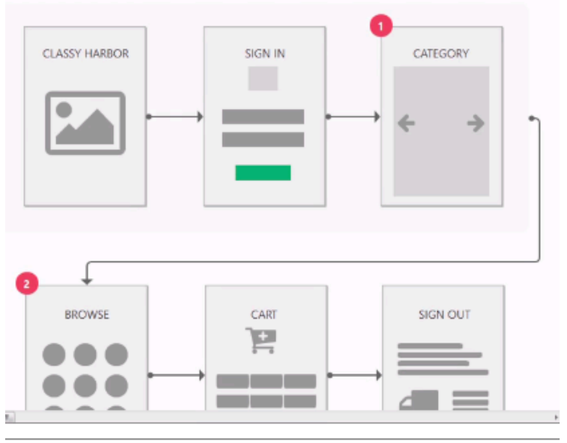
Some High Fidelity Prototyping Tools

- Marvel (https://www.youtube.com/watch?v=1439vIj-yDk)
- Balsamiq (https://www.youtube.com/watch?v=HebTVoOWMRs)
- Adobe XD (https://www.youtube.com/watch?v=53qdI7CPNxM)
- Sketch (https://www.youtube.com/watch?v=qywB0JHQeC4)
- InVision (https://www.youtube.com/watch?v=v10t2azNaFs)
- PowerPoint or Keynote…
- So many more…

Some High Fidelity Prototyping Tools

- Tools differ substantially in aesthetic and specificity
- Choose the one most appropriate for the feedback you want
- Every designer and design team has different preferences and rationales





Switching to summative evaluation

Why and where to evaluate

- Why?
  - Feedback on design directions and ideas
  - Discover major issues
  - (Help to) resolve disagreements
- Where?
  - In laboratory (controlled)
  - In natural settings (uncontrolled)

Empirical evaluation

- Empirical: involves users
  - Usability testing
  - Field studies
  - Click-through studies

Why do small studies?

- Nielsen (as well as Virzi and Lewis but less famously) found that you get *most* benefits (and bugs) from the first ~5 users
- Small studies have:
  - Little cost
  - Limited effects on development timelines
  - Can be done early and often
- Does this mean you never have to test with more than 5 users?
- What might have changed since these seminal findings in the early 1990's?

Why do large studies?

- Complex systems, complex people
- The data are there and it’s relatively easy to grab (think A/B testing)
- Require strong, statistically significant combination (e.g., regulation, safety)
- When your boss makes you

Why do formal lab testing?

- Require a controlled environment
- Replicability is a necessity
- Context of use is straightforward to model
- Context of use MUST be modeled

Why do field testing?

- Control and replicability are less essential
- Context of use is variable, nuanced, or otherwise difficult to model
- Need to reach a lot of people who cannot reach the lab
  (particularly useful to do use remote testing for these people)

It depends

- The answer to:
  - How many people are enough
  - Where to test with them
  - How controlled to be
  - How many tests to run them through

Field testing

Field testing

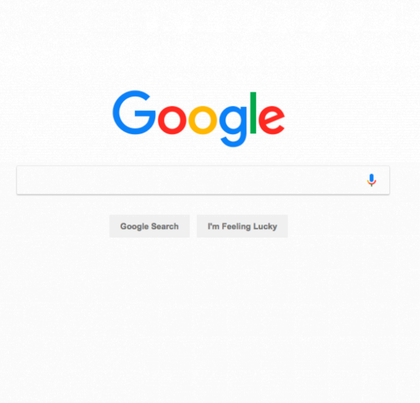
- Experiments in the lab are harder to generalize to new contexts
- Even harder to translate what we learn from a lab study to other tools/contexts
- Very hard in a controlled lab experiment to model the aggregate behaviors of groups (e.g., social media)

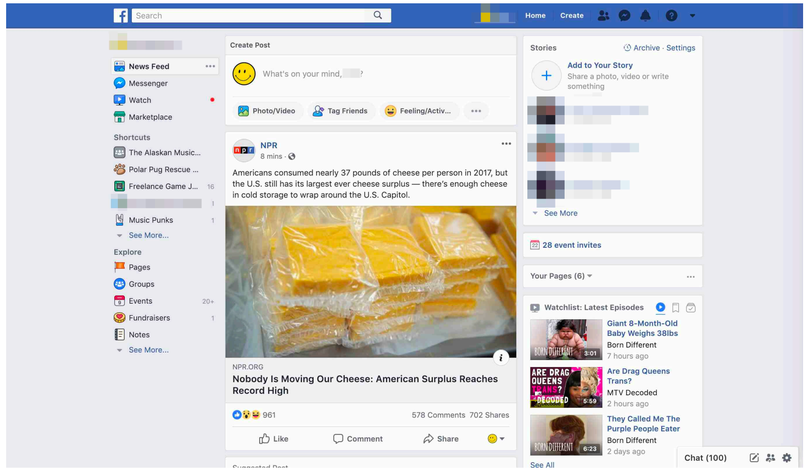
Living labs


Pinhole camera in wall
Camera (2 of 6) for people tracking
Everywhere Display Projector
Wireless PDA
Laser pointe tracker
Digital table
Projected image (can move anywhere)

Living labs

- Researchers inside every tech company (and many outside them) are doing A/B testing, log analysis, and generally experimenting on their users





TECHNOLOGY

Everything We Know About Facebook's Se Manipulation Experiment

It was probably legal. But was it ethical?

ROBINSON MEYER JUN 28, 2014

REUTERS

MORE

Why Rank High ALEXIS

Big T Expe Justi ALEXIS

Trun How News TAYLO

Maki Maki YouT TAYLO

Updated, 09/08/14

Facebook's News Feed—the main list of status updates, messages, and photos you see when you open Facebook on your computer or phone—is not a perfect mirror of the world.

But few users expect that Facebook would change their News Feed in order to manipulate their emotional state.

We now know that's exactly what happened two years ago. For one week in January 2012, data scientists skewed what almost 700,000 Facebook users saw when they logged into its service. Some people were shown content with a preponderance of happy and positive words; some were shown content analyzed as sadder than average. And when the week was over, these manipulated users were more likely to post either especially positive or negative words themselves.

This tinkering was just revealed as part of a new study, published in the prestigious *Proceedings of the National Academy of Sciences*. Many previous studies have used Facebook data to examine "emotional contagion," as this one did. This study is different because, while other studies have observed Facebook user data, this one set out to manipulate it.

The experiment is almost certainly legal. In the company's current terms of service, Facebook users relinquish the use of their data for "data analysis

Field studies

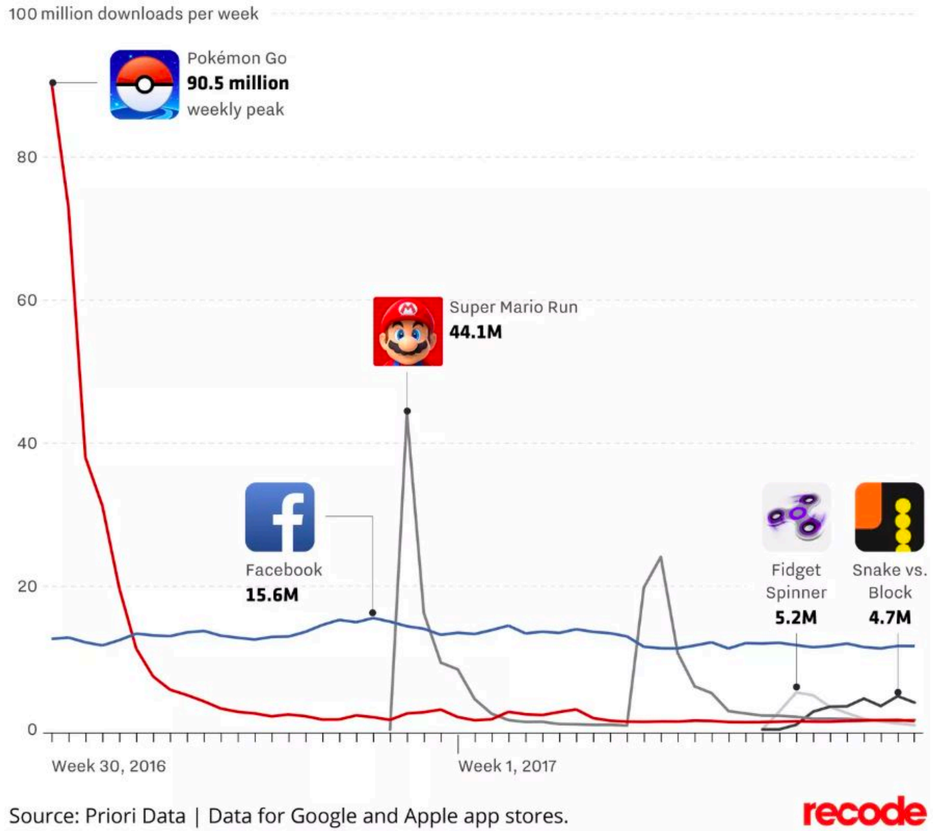
- Give people a functional prototype of your system
  and let them use it naturally for a set amount of time
  - Also called “in situ” studies, “real world” deployments, or “in the wild”

Considerations

- How long? / Novelty effect
- How many people?
- How to recruit?
- How to retain participants?
- Experiential or exploratory?
- What data to collect?

How long? / Novelty effect

- Any new technology will get the most use when it is first introduced, interest wanes
- How long?
  - Nathan Eagle (MIT) estimates 2 weeks
  - I think it's longer
- If the goal is to understand real-world use, field deployments should last long enough for the novelty effect to wear off



Participants

- The more the better (typically), but think about your resources
- Try to recruit a diverse a sample as possible (and as you need)
  - Think about recruiting proportional to your personas
  - If experimental, might want to recruit a homogenous sample to reduce variables
- Recruitment: internet ads, word of mouth, “snowball” sampling
  - Consider offering payment to attract and retain

Experimental or exploratory?

- Comparing new product against an old one can be very powerful
  - Your new product: experimental
  - Old product: control
  - “participants preferred using product X over product Y 9 times out of 10”
- Exploratory: just give out your product and see what happens
  - Often better in initial stage evaluation

What data to collect?

- Pre- and post- evaluation interviews and surveys
  - Depending on the length of the study, consider mid-study interviews as well
- Log usage data (if possible)
  - Automatically, with a computer script
  - Timestamped
- Diary entries after each use
  - Can automatically be prompted after usage

Today’s goals

By the end of today, you should be able to…

- Develop useful prototypes at higher fidelity than paper
- Think through decisions about study size and environment in summative evaluation
- Make choices about study length, style, and population in a field evaluation

IN4MATX 231:
User Interface Design & Evaluation

Class 12:
High-Fidelity Prototyping & Summative Evaluations

Daniel Epstein