



# User Research: Informing Product Decisions with Customer Insights

**Julie Stanford**

Lecturer, Computer Science Department  
Stanford Center for Professional Development

**This document is for your personal,  
educational, non-commercial use only.**

Copyright © Stanford University. All Rights Reserved.

# Who is Julie Stanford?

- **Science and Engineering Education Fellow, HCI Lecturer,**  
Stanford Computer Science (13 years)
- **Principal and Founder, Sliced Bread Design** (20 years)



Select a move type

Government moves the big stuff and you move the rest (HMO + PPM)

You move all your stuff (PM)

Government moves all of your stuff (MM)

**Pros and Cons**

**Pros**

- The government can arrange a mover to handle the move.
- Potential for you to save a little money by moving some items yourself.
- Professional or experienced movers handle moving them with you.

**Cons**

- More things to keep track of.
- More prone to damage what you move from what the

**Pros and Cons**

**Pros**

- You choose how your stuff is transported.
- Moving company uses a small truck for moving.
- Flexible moving dates (during the week, the weekend, or across months).
- Can still hire moving company or use a pod.

**Cons**

- You have to arrange everything.

**Pros and Cons**

**Pros**

- The government arranges moving companies to pack & transport your stuff.
- Less hassle.
- The claims process is available to you if anything becomes damaged or broken.

**Cons**

- Limited availability.
- Can only move on weekdays.

Source: US General Govt

# Course Overview

## 1: User Research Methods

- Overview
- Introduction to User Research
- Types of Methods and their Attributes
- User Research within the Product Development Process

## 2: Deep Dive: Rapid Experimentation

- What is Rapid Experimentation?
- How to Run a Rapid Experiment
- Subject Recruiting
- Synthesizing and Sharing Results

## 3: Deep Dive: Moderated User Testing

- What is Moderated User Testing?
- Test Moderation
- Creating a Moderated Test Plan
- Synthesizing and Sharing Results

## 4: Deep Dive: Unmoderated User Testing

- What is Unmoderated User Testing?
- Creating an Unmoderated Test Plan
- Synthesizing and Sharing results

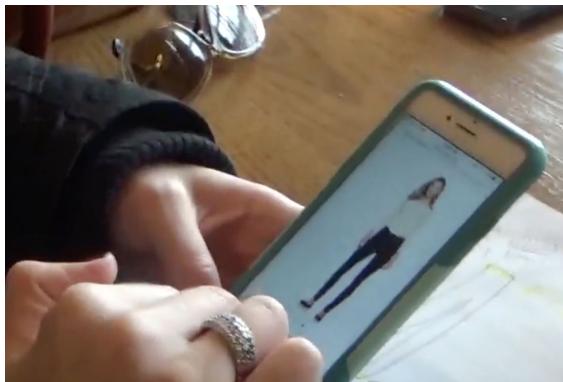
# Introduction to User Research

# What is user research?

Bottom line: Engaging with users to understand their point of view



# Example: Omni-channel denim shopping



# User research can help you...

- Gain empathy for your customer
- Understand customer's current behaviors, motivations and needs
- Gain insight into what's broken about current experiences
- Generate new innovative ideas you might not have considered
- Validate assumptions
- Discover how customers engage with your product in the real world
- Understand why things are happening
- Distinguish what customers say they are doing from what they are actually doing



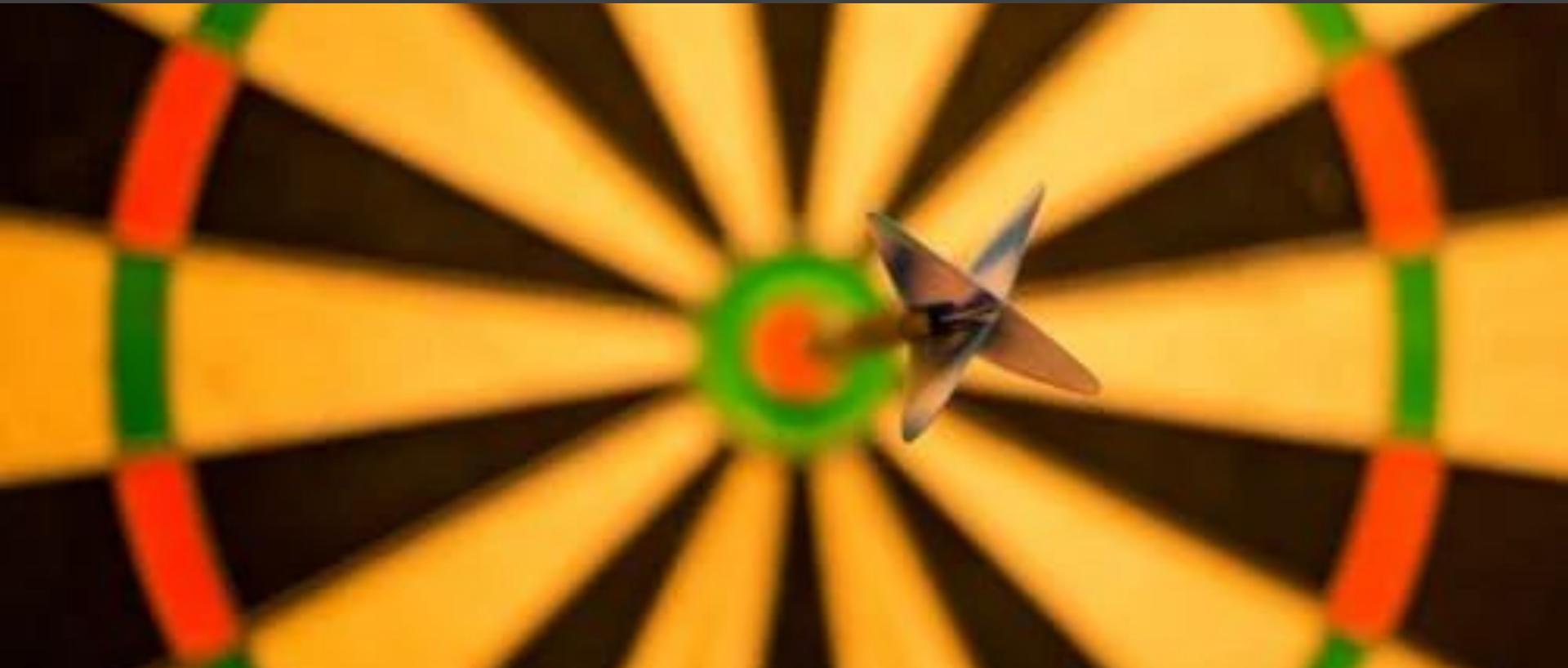
# What user research *can't* do



- Tell you exactly what to design
- Come up with new ideas for you
- Make decisions for you
- Predict the future

# **Product management end goal**

Create something customers keep coming back to

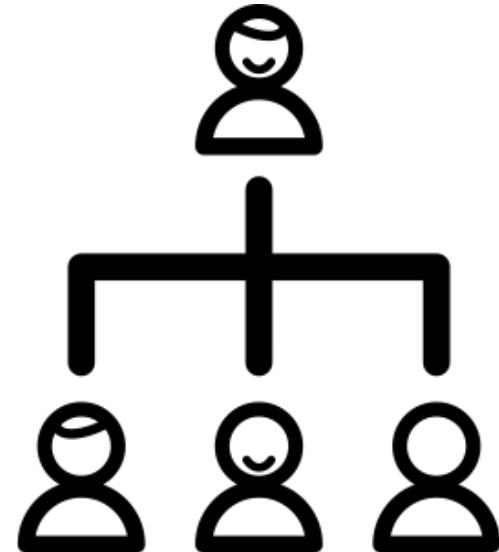


# Who does user research?

**Small orgs:** PMs conduct user research

**Larger orgs:** PMs collaborate with user research professionals:

- User researcher
- UX designer
- Information architect, etc...



Ideally involves **all stakeholders**

“We [can’t afford / don’t have time / don’t know how to do] user research.”

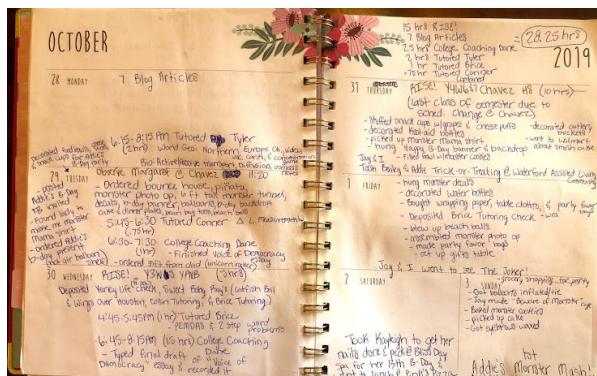


# Product Manager: attitude of biggest fan



# Types of Methods and their Attributes

# User research: What comes to mind?



What are the top 3 factors your employer uses to evaluate your job performance?

Fuel consumption  
 Ability to keep down costs  
 Courtesy and professionalism  
 Availability to take jobs  
 Safety record  
 Teamwork  
 Delivering or picking up on time  
 Miles driven per day  
 Other (please specify) \_\_\_\_\_

[PREV](#) [NEXT](#)

# Wide range of methods



Usability testing



Unmoderated user testing



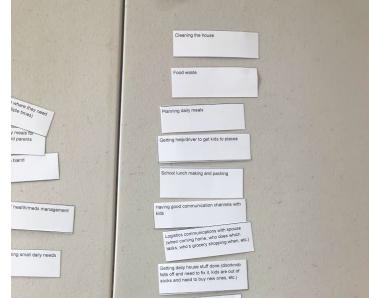
Rapid experimentation



Interviews



Contextual inquiry



Card sorts



Analytics & A/B Testing

What are the top 3 factors your employer uses to evaluate your job performance?

- Fuel consumption
- Ability to keep down costs
- Courtesy and professionalism
- Availability to take jobs
- Safety record
- Teamwork
- Delivering or picking up on time
- Miles driven per day
- Other (please specify) \_\_\_\_\_

Surveys

# Methods have key attributes

**Generative:** New idea generation

**Evaluative:** Evaluating ideas and designs

---

**Qualitative:** Small sample size, deep insight, exploration

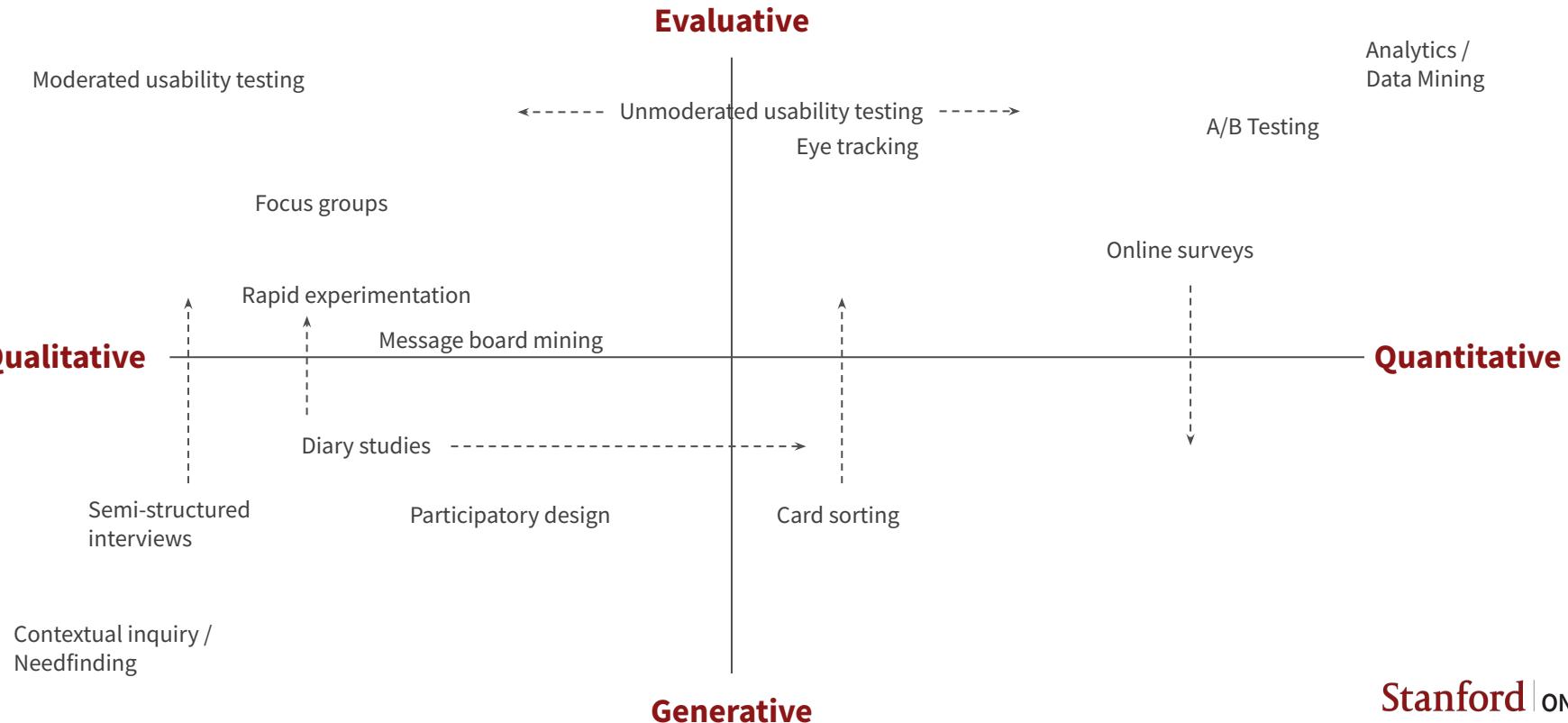
**Quantitative:** Large sample size, measurable, semi-objective insight

---

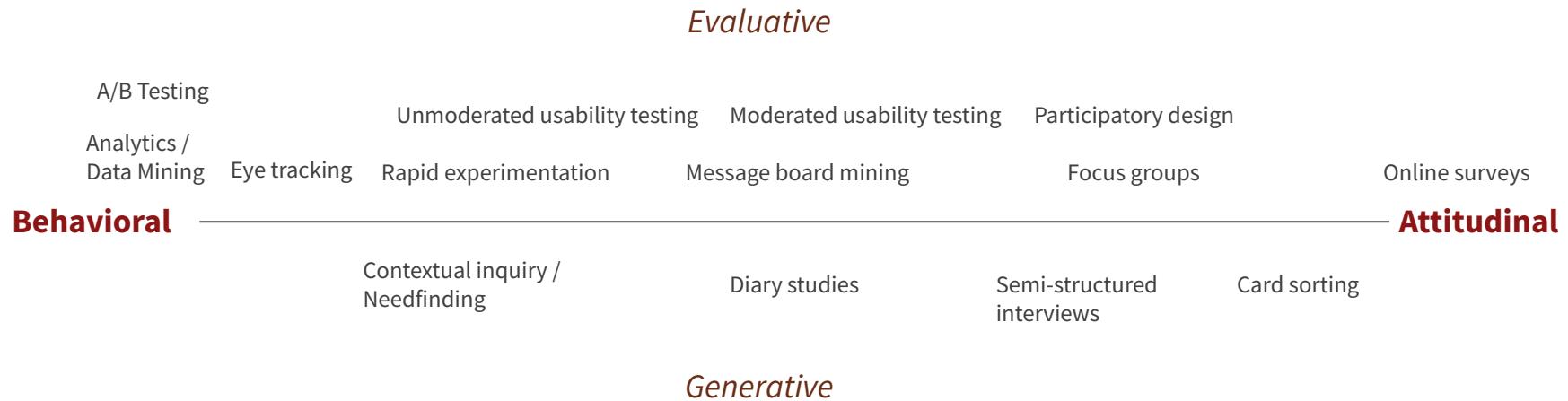
**Behavioral:** Observe what people do

**Attitudinal:** Listen to what people say

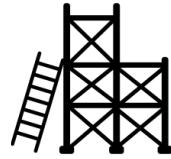
# Key attributes mapped



# Behavioral vs. attitudinal



# Methods have varying levels of structure



## Semi-structured

- Semi-structured interviews
- Moderating usability testing
- Participatory design
- Focus groups



## Experientially structured

- Contextual inquiry / Needfinding
- Rapid experimentation
- Diary studies



## Very structured

- Unmoderated usability testing
- Card sorting
- Surveys (online and via structured interviews)
- Analytics and A/B Testing
- Eye tracking
- Message board mining
- Focus groups

# Different flavors within methods

- Moderated vs. unmoderated
- Remote vs. in person
- Pre-recruited vs. Guerilla subject recruiting



# Our focus: 3 Key Methods for Evaluative Research

Rapid Experimentation | Moderated usability testing | Unmoderated testing

# Key Method 1: Rapid experimentation

Creating an experience using off the shelf pieces to observe how people would actually engage with your idea



## Attributes

- Qualitative
- Evaluative and generative
- Behavioral
- Experientially structured
- In-person or remote
- Moderated and unmoderated
- Pre-recruited or guerrilla

# Key Method 2: Moderated usability testing

Gathering feedback on design ideas by asking a user to imagine themselves in certain predefined scenarios facilitated by a moderator



## Attributes

- Qualitative
- Evaluative (mostly)
- Attitudinal and behavioral
- Semi-structured (depending on condition of prototype)
- In-person or remote
- Moderated
- Pre-recruited or Guerrilla

# Key Method 3: Unmoderated testing

Gathering feedback on design ideas by providing specific tasks for users facilitated by user testing software

## Attributes

- Qualitative or Quantitative
- Evaluative
- Behavioral and attitudinal
- Very structured
- Remote
- Unmoderated
- Pre-recruited

# Key method comparison table

	Qual/Quant	Evaluative vs. Generative	Behavioral vs. Attitudinal	Structure	Location	Moderation	Recruitment
Rapid Experimentation	Qual	Evaluative and Generative	Behavioral	Experientially structured	In-person or remote	Unmoderated and moderated	Pre-recruited or Guerrilla
Moderated Usability Testing	Qual	Evaluative	Behavioral and Attitudinal	Semi-structured	In-person or remote	Moderated	Pre-recruited or Guerrilla
Unmoderated Testing	Qual or Quant	Evaluative	Behavioral and Attitudinal	Very structured	Remote	Unmoderated	Pre-recruited

# Focus groups: Great for understanding groups



Lambert, Sylvie D., and Carmen G. Loiselle. "Combining Individual Interviews and Focus Groups to Enhance Data Richness." *Journal of Advanced Nursing* 62, no. 2 (April 2008): 228–37. <https://doi.org/10.1111/j.1365-2648.2007.04559.x>

Gill, P., K. Stewart, E. Treasure and B. Chadwick. "Methods of data collection in qualitative research: interviews and focus groups." *British Dental Journal* v204, no.6 (March 2008): 291 - 95. <https://www.nature.com/articles/bdj.2008.192>

# What about online surveys?

A structured questionnaire to understand and identify possible problem spots and high level attitudes OR validate assumptions with a large sample size

## Attributes

- Quantitative
- Evaluative
- Attitudinal
- Very structured
- Remote
- Unmoderated
- Pre-recruited or guerrilla

# Key method comparison table

	Qual/Quant	Evaluative vs. Generative	Behavioral vs. Attitudinal	Structure	Location	Moderation	Recruitment
Rapid Experimentation	Qual	Evaluative and Generative	Behavioral	Experientially structured	In-person or remote	Unmoderated and moderated	Pre-recruited or Guerrilla
Moderated Usability Testing	Qual	Evaluative	Behavioral and Attitudinal	Semi-structured	In-person or remote	Moderated	Pre-recruited or Guerrilla
Unmoderated Testing	Qual or Quant	Evaluative	Behavioral and Attitudinal	Very structured	Remote	Unmoderated	Pre-recruited
Online Surveys	Quant	Evaluative	Attitudinal	Very structured	Remote	Unmoderated	Pre-recruited or Guerrilla

**Surveys:** lack nuance and specificity; people are poor reporters

# How qual and quant work together



**Fire map**  
Quantitative studies, analytics and surveys identify **where** problems might be happening



Qual

**Investigation**  
Qualitative, open-ended research explores **why** these problems are happening and what underlying causes exist

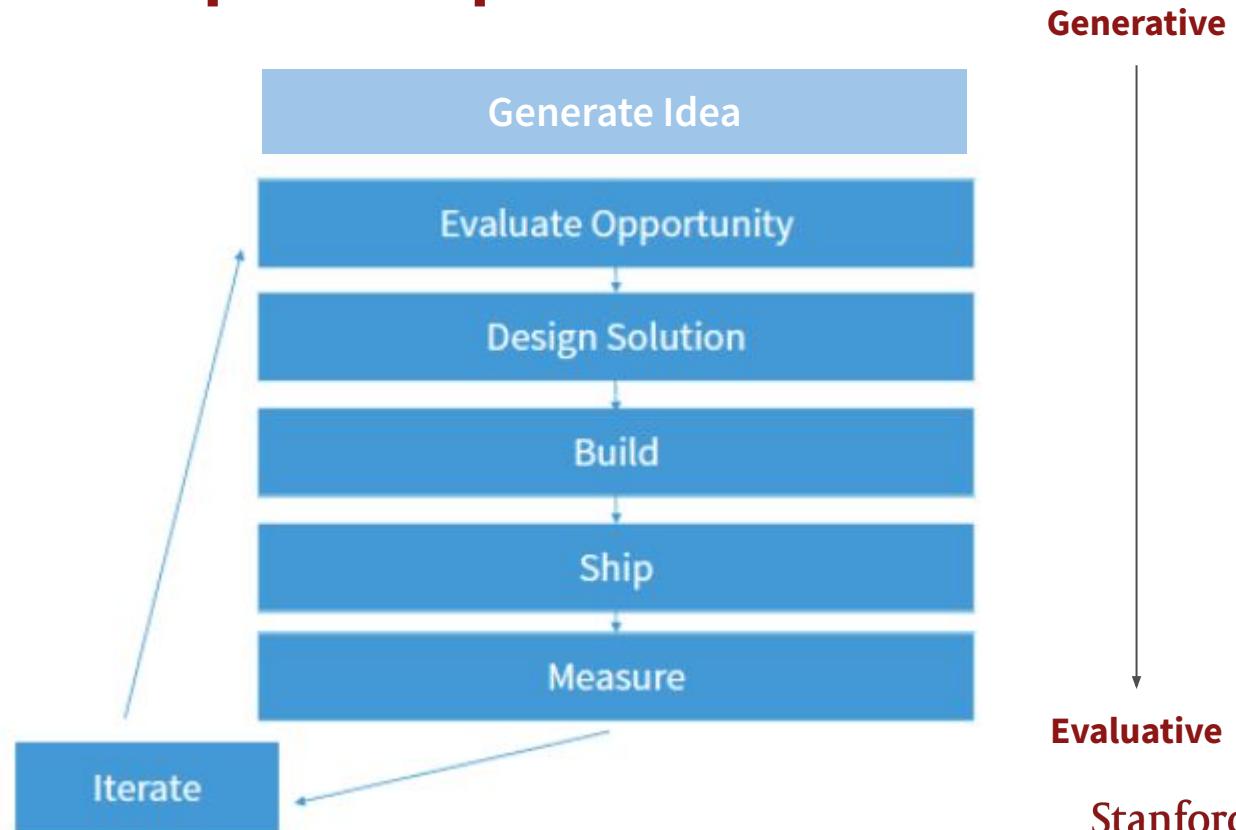


Quant

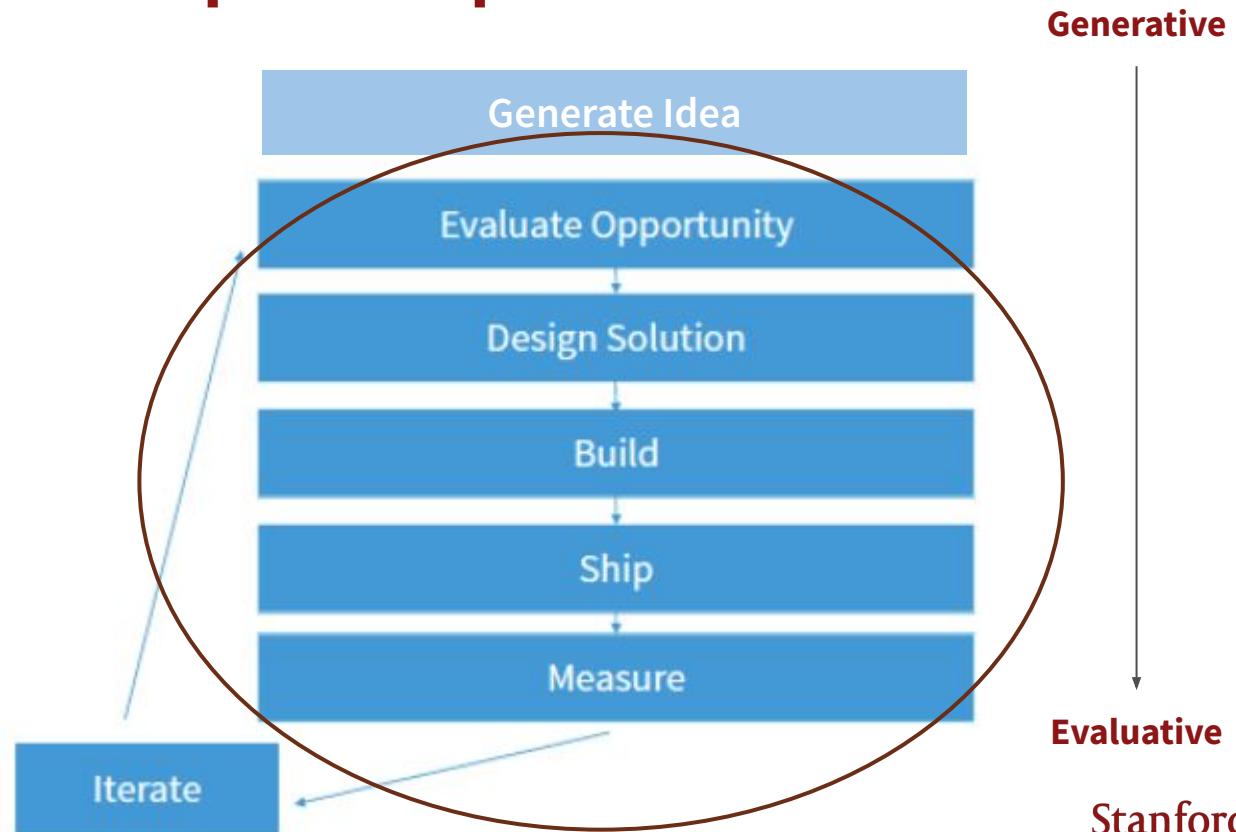
**Findings validation**  
Surveys can then **validate** key insights from a qualitative study with a larger sample size

# User Research within the Product Development Process

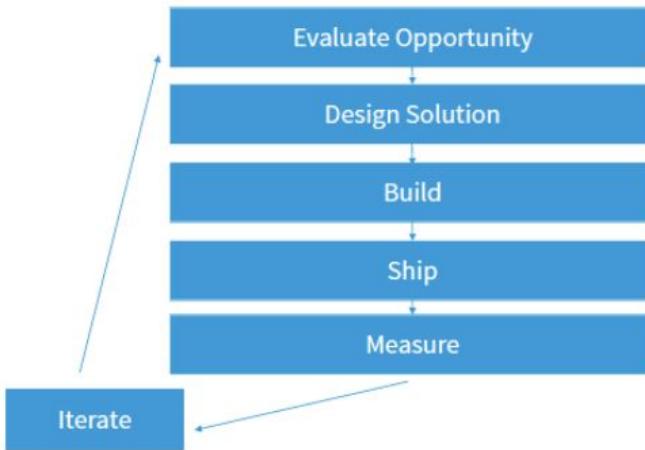
# Product development process



# Product development process



# Different research goals at different stages



- **Evaluate Opportunity:** Understand the market, assess match between the idea and user needs, determine what is relevant for users, refine requirements
- **Design Solution:** Evaluate designs, decide between different ideas
- **Build and Ship:** Refine final design, make sure there are no surprises at launch
- **Measure:** Evaluate acceptance, prepare for next version, find new problem areas, identify trends

# Evaluate opportunity

Understand the market, assess match between the idea and user needs, determine what is relevant for users, refine requirements

- What do people need? What are innovation opportunities?
- Will people use this product/feature? How will they use it?
- What features should we include?
- What needs to be designed?

## User research options

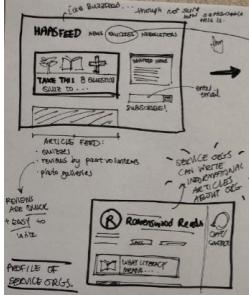
- Evaluative (Specific idea selected)
  - Rapid experimentation
  - Semi-structured interview with concepts
  - Online surveys (post qualitative research)
- Generative (Generating wide range of ideas)
  - Contextual inquiry / Needfinding
  - Diary study
  - Card sorting



Experiential prototypes



Concept sketches and storyboards



# Design solution

Evaluate designs, decide between different ideas

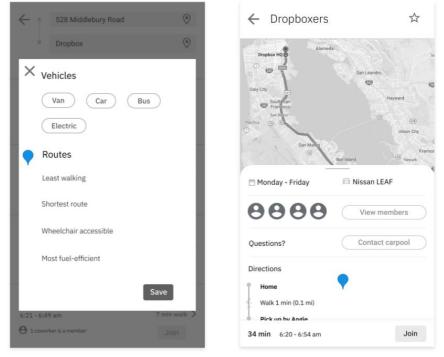
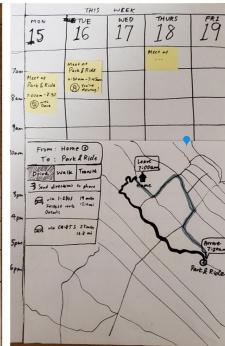
- What's the right way to design this experience?
- Is this design easy to use/ usable ?
- Which design approach is the best?
- How do we refine this design?

## User research options

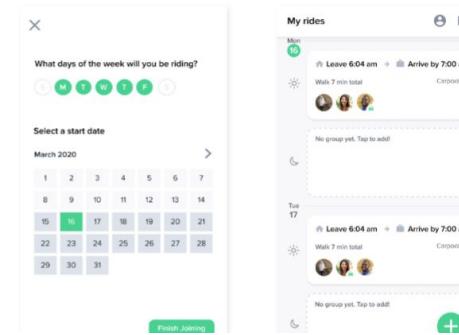
- Evaluative
  - Moderated usability testing
  - Unmoderated usability testing
  - Eye tracking
  - Card sorting
  - *Rapid experimentation*



Paper prototypes



Wireframes, medium fidelity designs

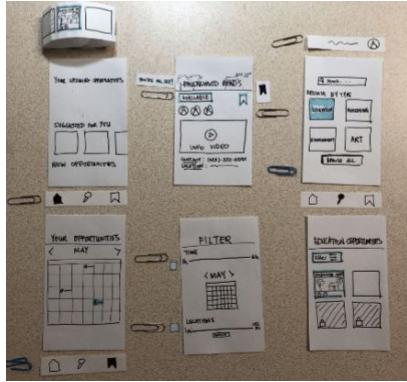


High fidelity designs, clickable prototypes

# Design solution

Evaluate designs, decide between different ideas

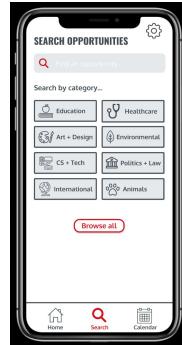
- What's the right way to design this experience?
- Is this design easy to use/ usable ?
- Which design approach is the best?
- How do we refine this design?



Paper prototypes



Wireframes, medium fidelity designs



High fidelity designs, clickable prototypes

## User research options

- Evaluative
  - Moderated usability testing
  - Unmoderated usability testing
  - Eye tracking
  - Card sorting
  - *Rapid experimentation*

# Build and ship

Refine final design and make sure there are no surprises at launch

- How are people using the actual product?
- What small tweaks have a strong impact?
- How should we launch this?
- What have we missed?

## User research options

- Evaluative
  - Unmoderated usability testing
  - Diary study
  - A/B testing
  - Data analytics
  - Eye tracking



A fully built, ready to launch experience.

# Measure

Evaluate acceptance, prepare for next version, find new problem areas, identify trends

- How is this being adopted?
- Is this product meeting its goals?
- What should we work on next?

## User research options

- Evaluative
  - Unmoderated usability testing
  - Message board mining
  - Data analytics
  - Eye tracking
  - Online surveys
- Generative
  - Semi-structured interviews + moderated usability testing
  - Diary study
  - Contextual inquiry / Needfinding



A fully launched experience ready for V2

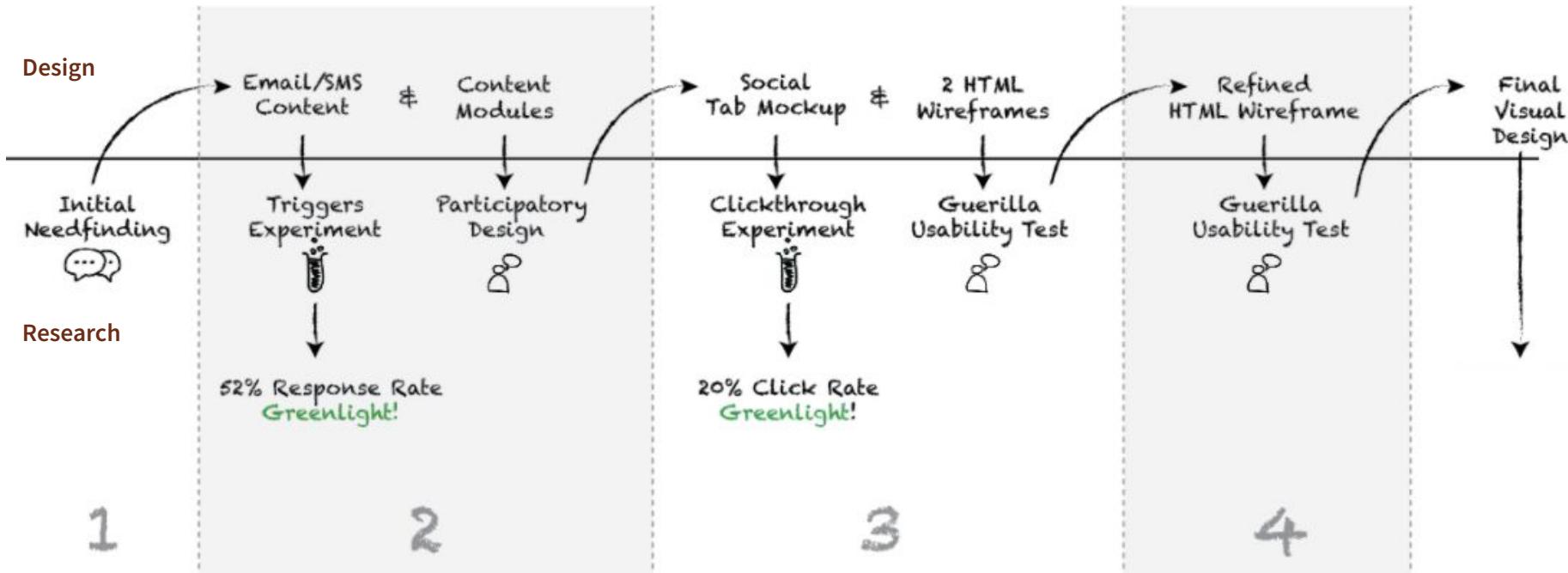
# Selecting a Method

# Case Study 1: Multiple Methods During One Project

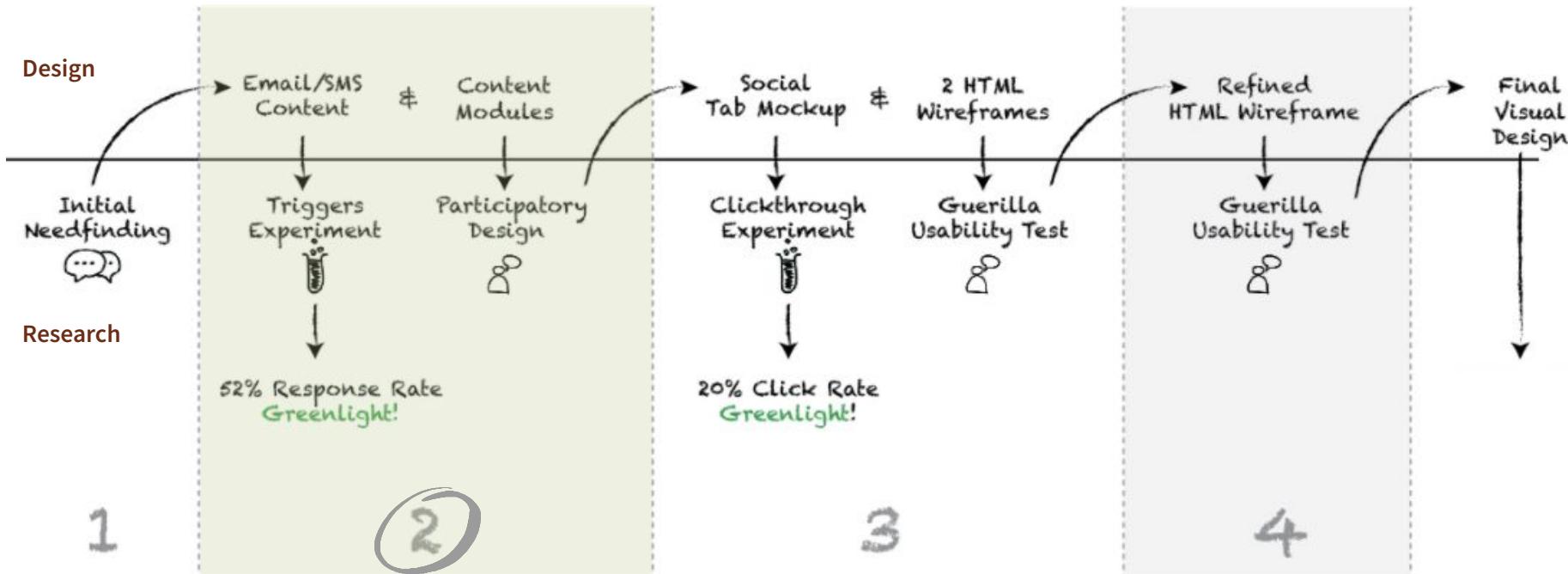
# Internal project: Employee directory

- Make a truly useful directory profile page
- Provide a single source for people to discover social content about another employee
- Make it easy to keep your information up to date in the directory

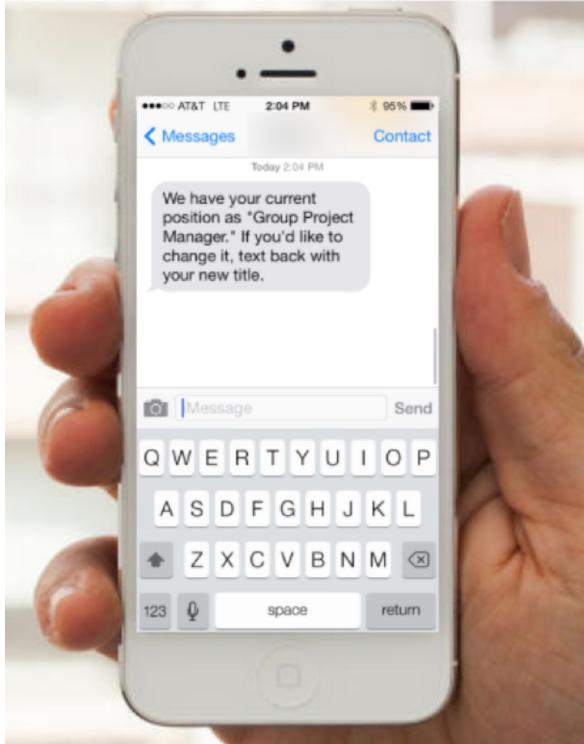
# Plan: Iterative user research and design



# Plan: Iterative user research and design



# Cycle 2 user research

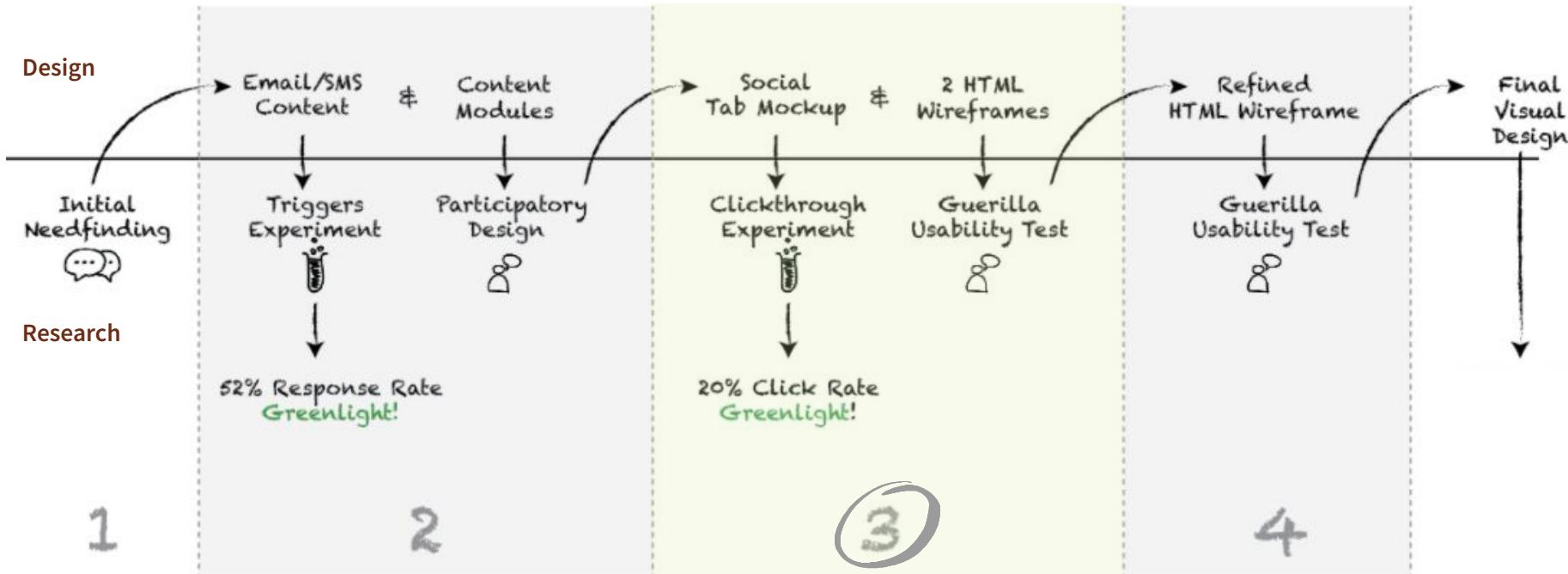


*SMS Rapid Experiment*



*Participatory design + semi-structured interviews*

# Plan: Iterative user research and design



# Cycle 3 user research

**Contact Info** **Social Info**

There are no tags yet for this profile.

**Shawn Wilson**  
Group Product Manager  
Employees  
650-944-0112  
[View Profile](#) [Edit Profile](#)

**Contact Information**

Name	Shawn Wilson
Office number	65112
Extension	65112
Mobile number	415-200-2712
Email	<a href="#">Shawn.Wilson@intuit.com</a>
Department	4402 - PR/HR/Comm
Company	Intuit Corp
Group	888 - Corp-Wdg
Site Name	Mountain View
Employee Type	Employee
Billing	\$0
Phone	650-944-0112
Code/Office #	00020112
Date hired	07/19/02
Street Address	2800 Cesar Chavez St, Mountain View, CA, 94037

**About Me**

**Background**  
I graduated from University of California at Santa Cruz with degrees in Politics and Economics and currently live with my wife, Kristen, in San Francisco.

**INSIGHT TOP NAV HERE**

**Susan Hwong**  
Group Product Manager, QuickBooks Payroll  
In Workday: Group Initiative Manager

**Email** [susan\\_hwong@intuit.com](#)  
**Work Phone** 650-944-5555  
**Mobile Phone** 415-555-5555  
**Location** Mountain View, building 7, floor 2, office 0709100

Photos 14

**Org Structure** Manager: [Pratiksha Singhal](#) Direct Reports: 6 **About Me** 7 years at Intuit **Social Info** Last Yammer: 09/18

**INSIGHT TOP NAV HERE**

**Susan Hwong**  
Group Product Manager, QuickBooks Payroll (Group Initiative Manager)  
[susan\\_hwong@intuit.com](#)  
**Work Phone** 650-944-5555  
**Mobile Phone** 415-555-5555  
**Fax** 650-555-5555  
**Yammer** [susan\\_hwong](#)

Photos 14

At Intuit 6 years

Mountain View Time 3:48PM  
Bldg 7  
Floor 2  
Office 0709100  
Address v [more Intuit roles v](#)

[My LinkedIn Profile](#)

**About Me**

What teams have you worked on at Intuit?  
[QuickBooks Payroll](#) [QuickBooks](#) [QuickBooks Online Edition](#)

What are your skills and expertise?  
Product Management | Go-to-market Strategy | Agile Methodologies | SaaS | Product Marketing | Analytics | Product Launch | Product Requirements | User Experience

What tools and technologies are you involved with or using?  
{tool icons here}

What previous companies have you worked for?  
[eBay](#) (2000-2002, Product Manager)

**Organization**

**My Team**

**My Department** QuickBooks Pro and Premier  
**My Business Unit** QuickBooks

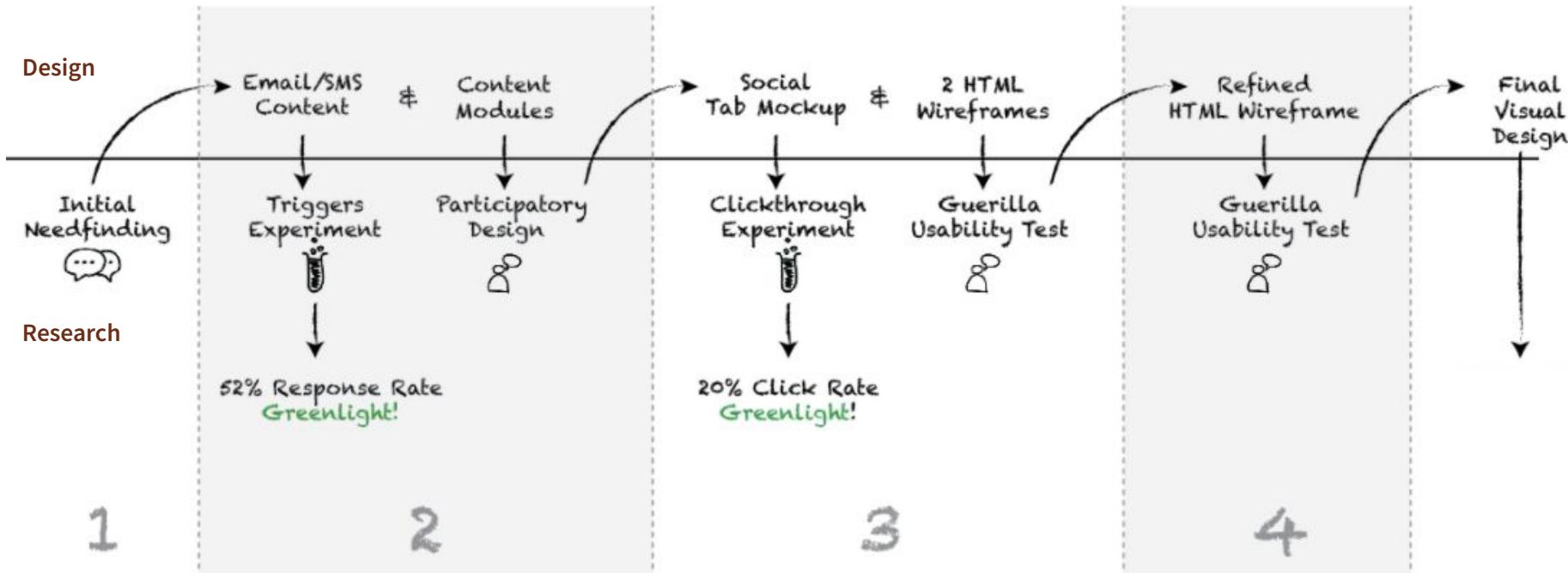
**Org Structure** Org structure goes here

Rapid experiment: Social info tab

Moderated user study: A/B wireframe concepts

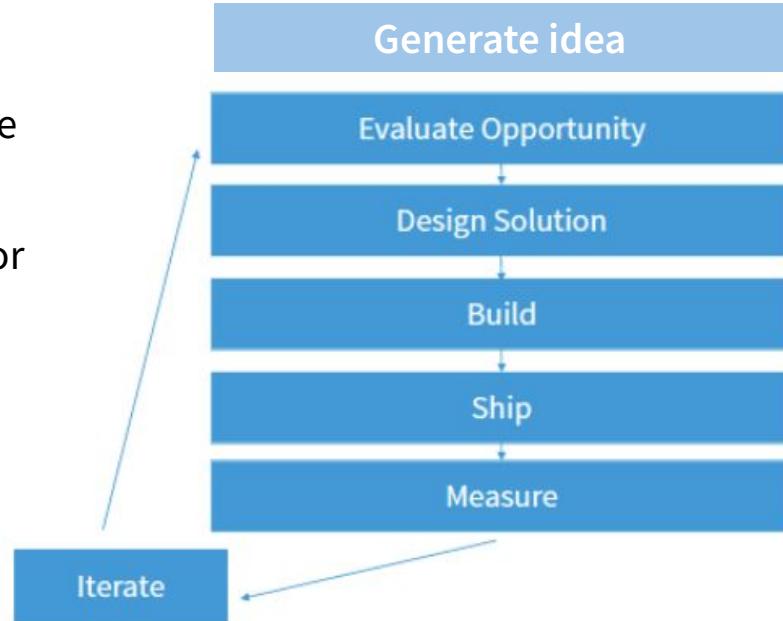
Stanford | ONLINE

# Plan: Iterative user research and design



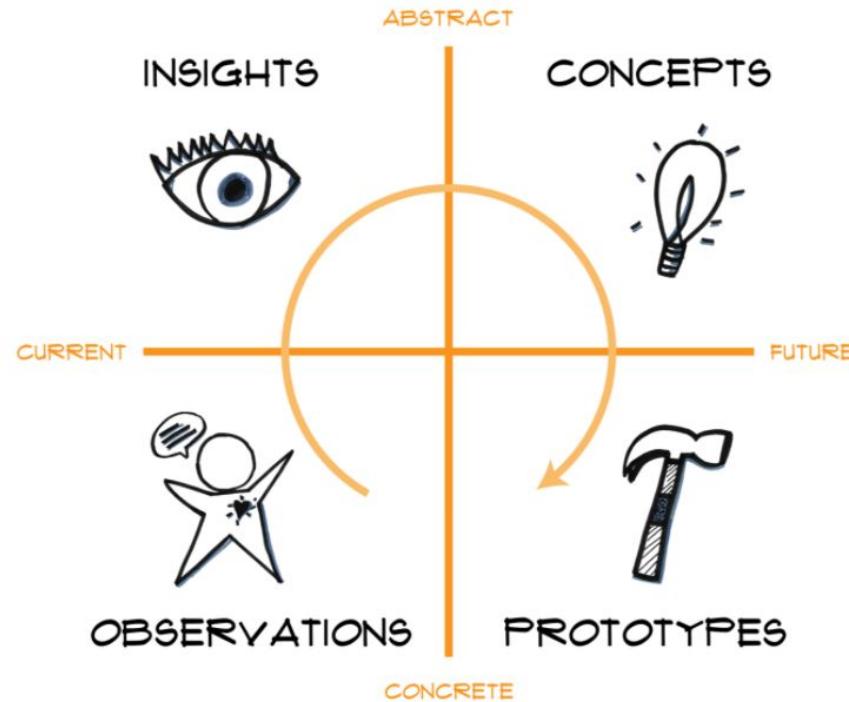
# Method selection criteria

- PD phase
- **Questions you are trying to answer**
- Types of prototyping options available
  - Experience, paper, mid to high fidelity prototypes, or actual implementation
- Time and resources available
  - 1 hour vs. 1 week vs. 1 month

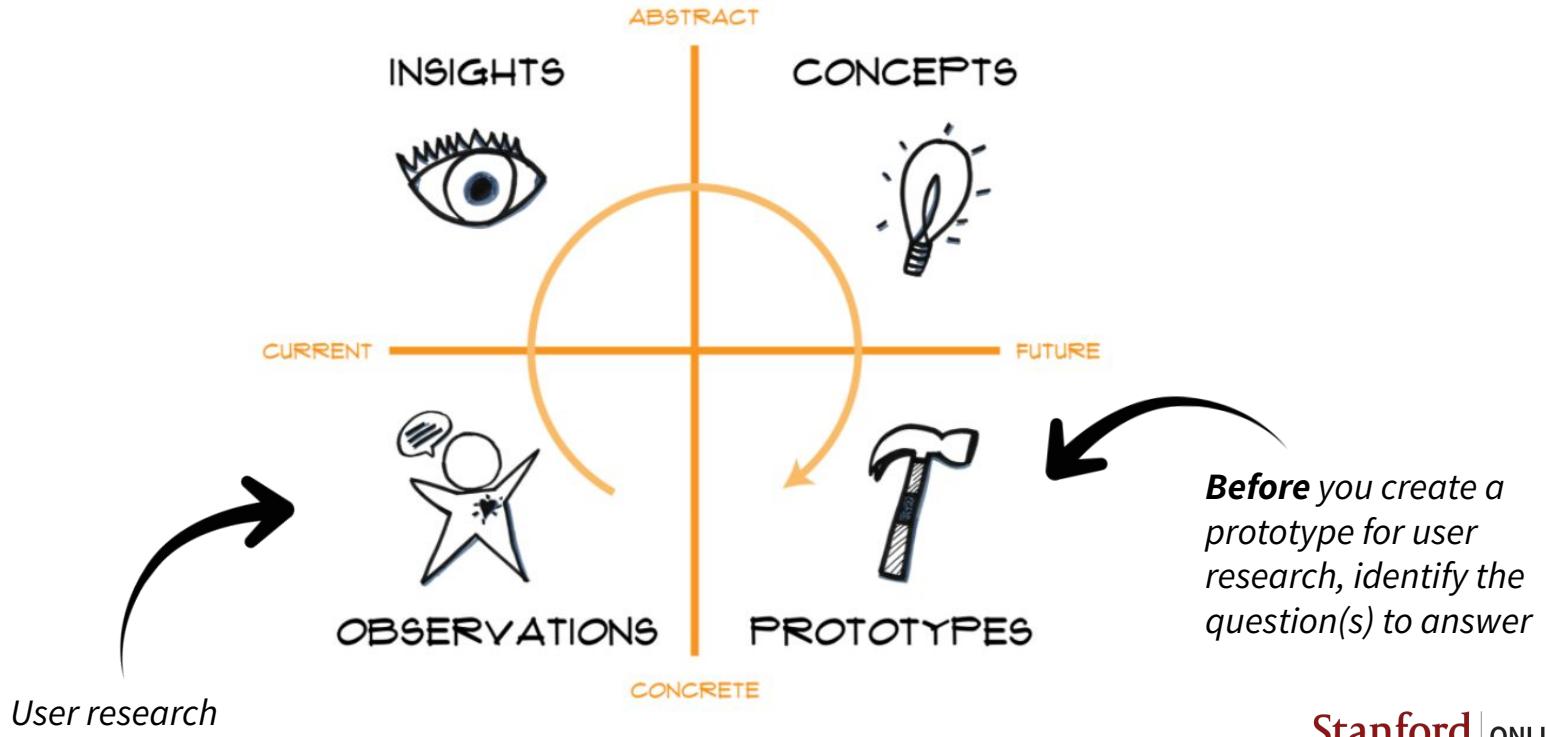


# Questions to Answer

# Design process overview



# Design process overview

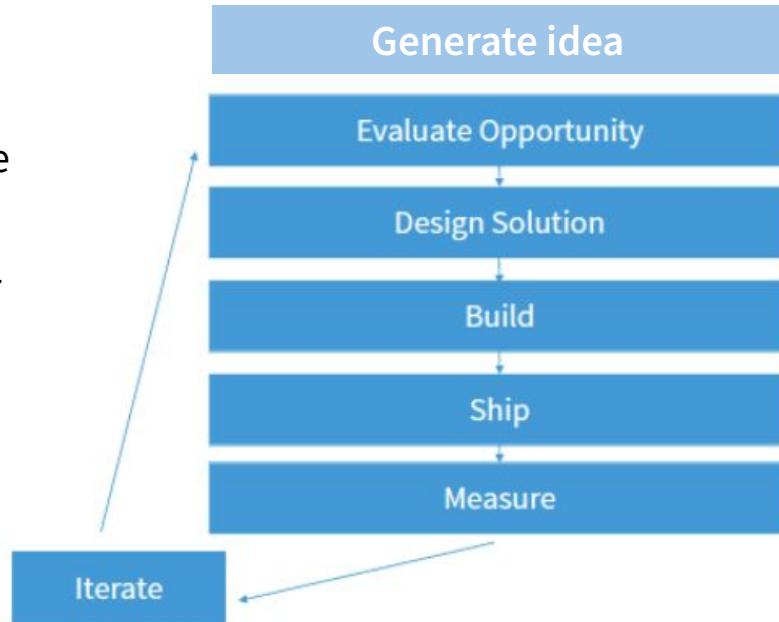


# Make a list of questions and rank them

- Are people interested in feature X more than feature Y?
- Will people actually use feature X?
- Where should this button go?
- Do people understand the language we are using to describe XYZ?
- What design will lead to more people doing X?
- How will the new design effect previous features ABC?

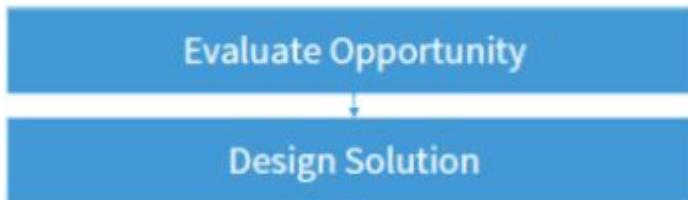
# Method selection criteria

- **PD phase**
- **Questions are you trying to answer**
- Types of prototyping options available
  - Experience, paper, mid to high fidelity prototype, or actual implementation
- Time and resources available
  - 1 hour vs. 1 week vs. 1 month



# Key Method 1: Rapid experimentation

Creating an experience using off the shelf pieces to observe how people would actually engage with your idea



## Key Questions: What are we designing?

- Will people actually use this feature / idea?
- How will people use this?
- Which features should we include?
- What needs to be designed?

## Type of prototype

- Experiential concept made from off the shelf components, not a test of the user interface

## Time and resources

- 2 hours quick and informal
- 1-2 weeks+ for use over time

# Key Method 2: Moderated usability testing

Gathering feedback on design ideas by asking a user to imagine themselves in certain predefined scenarios facilitated by a moderator

*Can be paired with participatory design, card sort and eye tracking*

Design Solution

**Key questions: How should this be designed?**

- What other features should we include?
- What's the right way to design this?
- Is this design easy to use / usable?
- Which design approach is best?
- What are the tradeoffs between options?
- How do we refine this design?

**Type of prototype**

- Any kind of prototype - low to high fidelity
- Implemented build

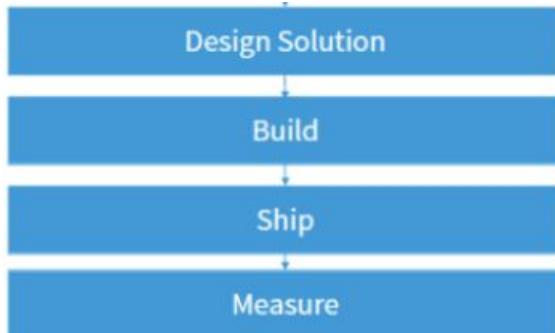
**Time and resources**

- 2 hours quick and informal
- 1 day to a week if more formal

# Key Method 3: Unmoderated testing

Gathering feedback on design ideas by providing specific tasks for users facilitated by user testing software

*Can be paired with card sort*



**Key questions: How should this be refined?**

- Is this design easy to use / usable?
- Which design approach is best?
- How do we refine this design?
- What small tweaks have a strong impact?
- What have we missed?
- How are people using the actual product?

**Type of prototype**

- Standalone mid to high fidelity mockups
- Implemented build

**Time and resources**

- Depends on what users need to be recruited
- Can create a quantitative test which may help certain stakeholders

# Moderated vs. unmoderated testing

## Moderated

- Can accommodate any type of prototype for testing
- Moderator can adjust as needed, improvise and ask follow-up questions
- Can provide answers about not just what, but why
- Can inject bias of moderator or accidental disclosure of correct use
- Time-consuming because moderator must be present
- Qualitative (except in rare cases)

## Unmoderated

- Test participants “moderate” themselves, lower overhead
- Test participants can test anytime, on their own, in a natural environment
- You can watch videos whenever you are free, not when subjects are available
- Offers quantitative option -- although this can be time consuming/expensive
- No way to intervene if something goes wrong or ask follow-ups
- Hard to get to bottom of “why”
- Must provide structured prototype -- no room for open ended-exploration for low fidelity mockups.  
Great for single page tests or wizards.

# Qualitative vs. quantitative user testing

3-8 subjects

30+  
subjects

## Qualitative testing

- 5 people will find the majority of issues\*
- Smaller issues will be masked by the more egregious ones. You need to fix them before you do more testing.
- No need to have quantitative results to see there is a problem in most cases. It is really obvious.
- Better to do a larger number of smaller tests than one large one.

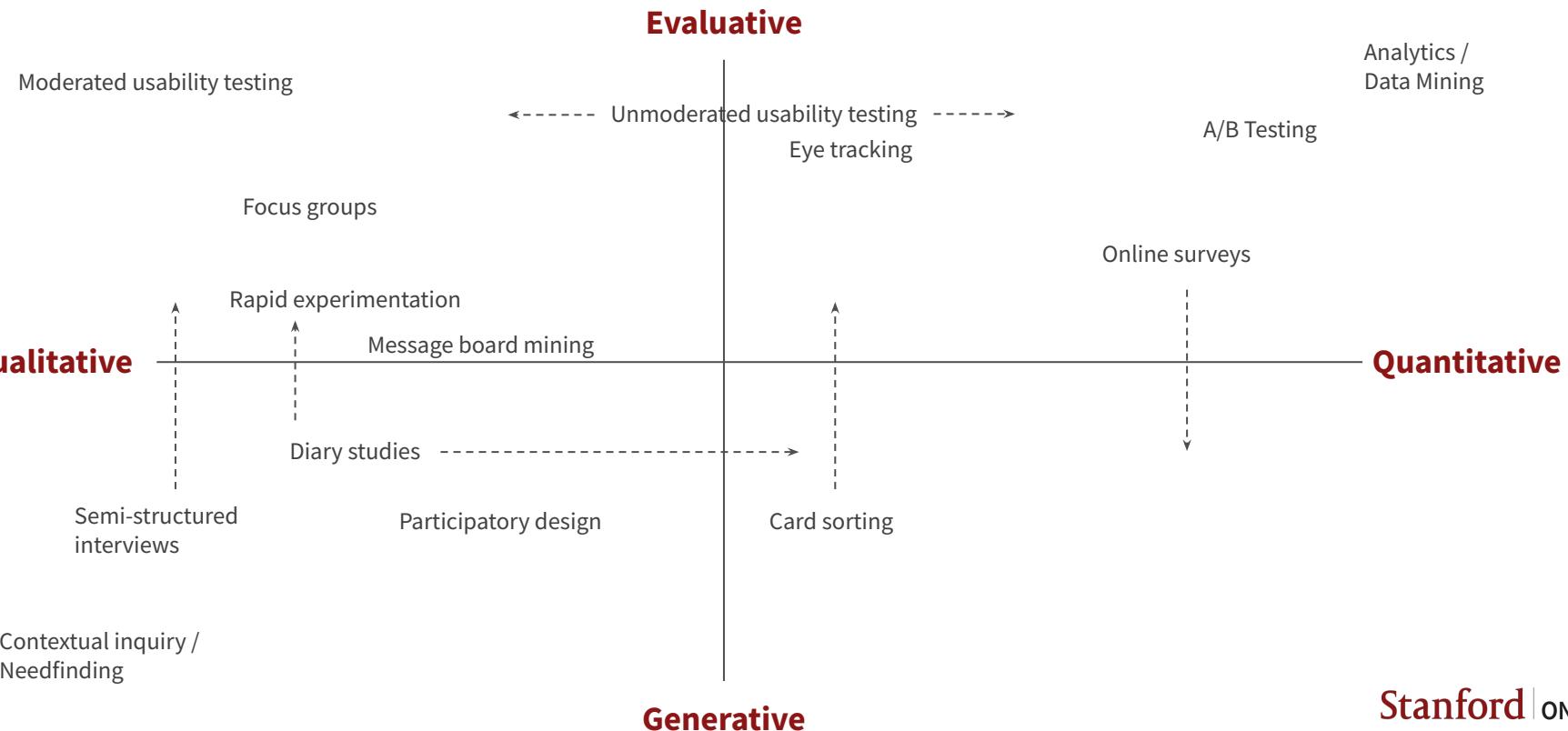
## Quantitative

- Useful for specific studies like card sorting or eye tracking.
- Can help for high stakes decision making for quant focused stakeholders.
- May need more subjects if backgrounds and points of view differ greatly -- for example in different countries.
- Can be expensive and often unnecessary.

# Methods table revisited

	Questions answered	Prototype type	Time and resources	Stakeholder needs
<b>Rapid Experimentation</b>	What are we designing?	Experiential	2 hours to 1-2 weeks	Qual only
<b>Moderated Usability Testing</b>	How are we designing it?	Any kind of prototype - low to high fidelity or implementation	2 hours to 1 week	Qual only
<b>Unmoderated Testing</b>	How should this be refined?	Standalone mid to high fidelity mockups or implementation	Depends on users that need to be recruited	Qual or quant

# Map of methods



**What method is best?  
The one you actually do.**

# Case study: Multiple methods of evaluation

Evaluate opportunity

Design solution

Build and ship

Measure

Semi-structured interview  
with paper prototype

Regular moderated on  
designs in increasing  
levels of fidelity

Data analytics for beta  
designs launched in A/B  
bucket testing

Data analytics for  
launched product

Rapid experiment on live site  
with 1 day mockup

Unmoderated test of near  
final designs

Unmoderated testing of  
new designs on live site

Unmoderated testing of  
design on live site versus  
competitor sites

You try it:  
What method would you choose?

# Example 1: What would you do?

You are a project manager for an online web dating app, and you've been asked to work on a new idea for a **remote video speed-dating feature** the PD team thinks might meet the market need during the COVID-19 pandemic.

You've been asked to explore the potential of this idea further to assess its efficacy and impact on user engagement.

Unfortunately, you have **no prototypes on hand** right now except for a few rough sketches as this scenario is unprecedented, but on the other hand, anything is possible! Since you are not sure yet if this proposal will be profitable or successful, you are **limited on the financial resources** you can use, but you have **plenty of time** to conduct your research and present your findings in your team's bi-annual feature proposal showcase in a month.

# Recommendation: Rapid experimentation

Creating an experience using off the shelf pieces to observe how people would actually engage with your idea



## Key Questions: What are we designing?

- Will people actually use this feature / idea?
- How will people use this?
- Which features should we include?
- What needs to be designed?

## Type of prototype

- Experiential concept made from off the shelf components, not a test of the user interface

## Time and resources

- 2 hours quick and informal
- 1-2 weeks+ for use over time

## Example 2: What would you do?

You work closely with the program management team on the implementation of a **song recommendation system for a music streaming app**, a feature they are calling "enhance my playlist".

The team already has **various medium-fi prototypes** of different implementation options for where the buttons, icons, and other components involved in the song recommendation user flow should go.

However, they are **unsure if users will resonate with the phrase "enhance" for recommendations** and want to get some insight on which options for different designs make the most sense. As this is a medium sized company, they are willing to support you with any resources you might need but would prefer to **keep expenses to a minimum** and get some **feedback within a week**.

# Recommendation: Moderated usability testing

Gathering feedback on design ideas by asking a user to imagine themselves in certain predefined scenarios facilitated by a moderator

**Key questions: How should this be designed?**

- What other features should we include?
- What's the right way to design this?
- Is this design easy to use / usable?
- Which design approach is best?
- What are the tradeoffs between options?
- How do we refine this design?

**Type of prototype**

- Any kind of prototype - low to high fidelity
- Implemented build

**Time and resources**

- 2 hours quick and informal
- 1 day to a week if more formal

Design Solution

# Example 3: What would you do?

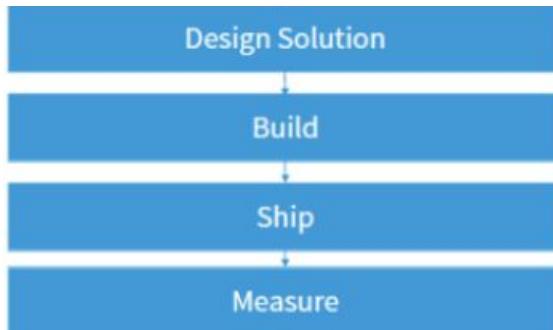
You are a program manager who works closely with the engineering team on a **smart fridge hardware company** that lets you monitor and control your fridge from a mobile app.

The engineering team is in the final phases of their hardware development cycle, and their engineering manager asked you to do a **final review of the UX before the product is launched in two weeks to catch last minute bugs, problems, or adjustments that need to be made**. Since it is nearing the official launch, you have production ready hardware and software models to conduct your research.

**Turnaround time must be quick** and the engineering team is expecting both qualitative insights and quantitative measurements in case anything needs to be changed. They've **offered to pay for any online services you might need** to conduct your research.

# Recommendation: Unmoderated testing

Gathering feedback on design ideas by providing specific, tasks for users facilitated by user testing software



**Key questions: How should this be refined?**

- Is this design easy to use / usable?
- Which design approach is best?
- How do we refine this design?
- What small tweaks have a strong impact?
- What have we missed?
- How are people using the actual product?

**Type of prototype**

- Standalone mid to high fidelity mockups
- Implemented build

**Time and resources**

- Depends on what users need to be recruited
- Can create a quantitative test which may help certain stakeholders

# Key Takeaways

1. User research is a great tool for **understanding your customers** and **gathering their feedback on your ideas** throughout the product development process
2. There are a **wide range** of user research methods. Methods have the following attributes:
  - good for **generating** vs. **evaluating** new ideas
  - **qualitative** vs. **quantitative**
  - **behavioral** (observe what people do) vs. **attitudinal** (listen to what people say)
  - **unstructured** vs. **semi or very structured**
3. Three key, effective methods that are widely used are: **rapid experimentation**, **moderated user testing** and **unmoderated user testing**
4. Choosing which method to use depends on your **stage in the product development** process, the **questions** you are answering, the **state of your prototype** and how much **time** and resources you have available



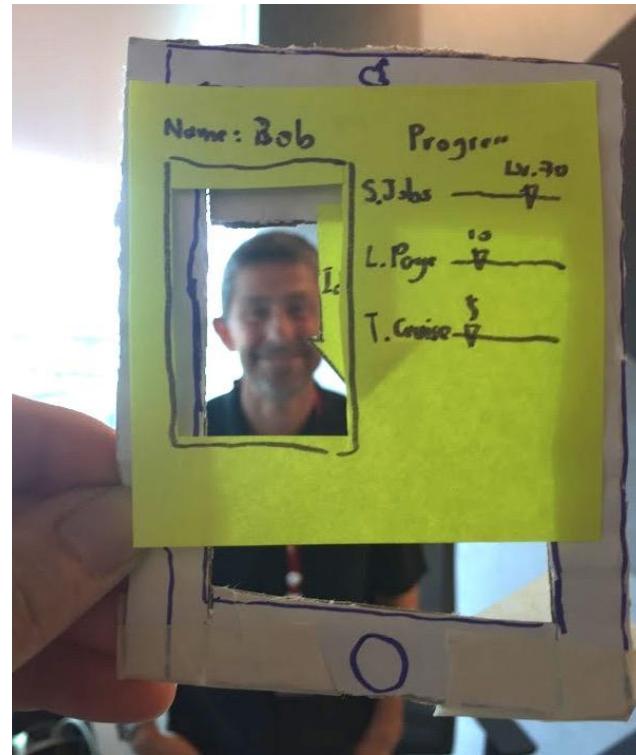
# Part 2: Deep Dive Rapid Experimentation

Stanford Center for Professional  
Development

# What is rapid experimentation?

**Generating experiments that involve prototypes to create an experience that evokes behaviors to validate if our design meets users' needs**





# Example: Parents, kids & social media

Idea:

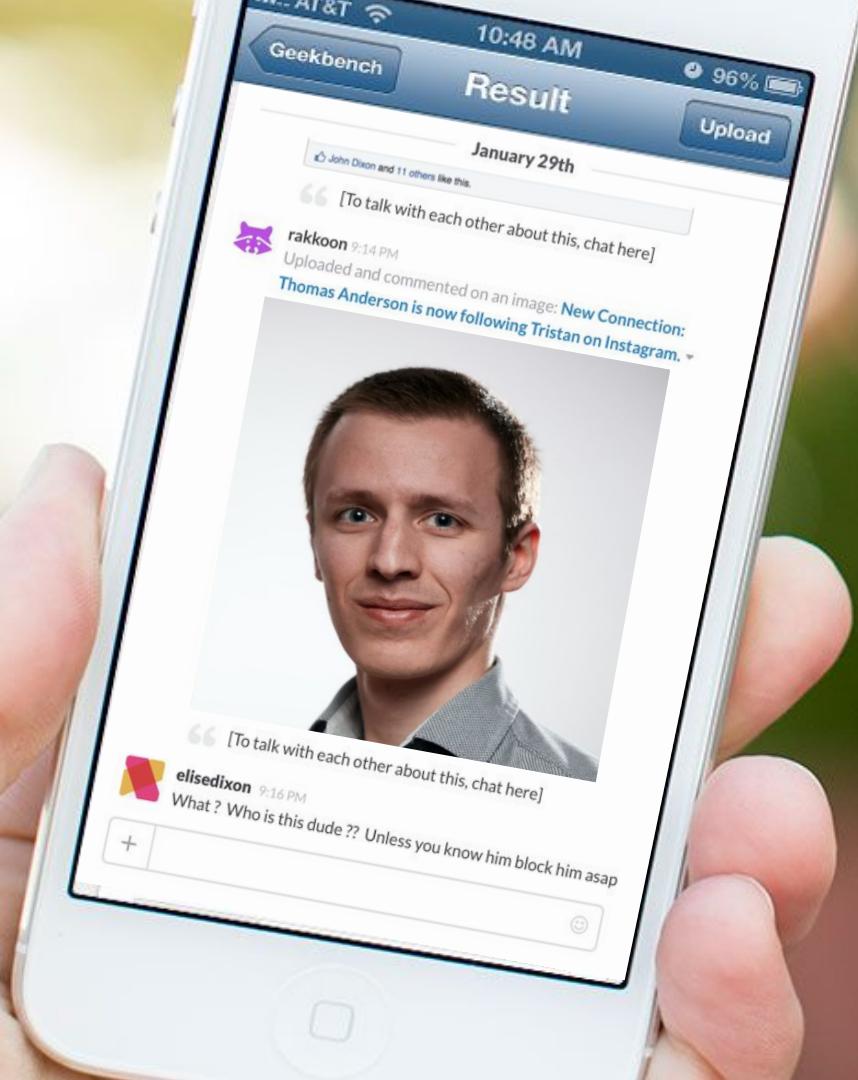
AI system that sends parents AND kids social media alert messages to improve communication by facilitating discussion

# Example: parents, kids & social media

Question:

Will parents and kids have conversations when they are sent social media alerts about possibly unwanted social media posts in a digital discussion space?

“Welcome!  
We’ll be using  
Slack to send  
alerts about  
Sam’s social  
media use. You  
can chat with  
each other about  
the alerts right in  
this app.”



**rapid experiments separate  
what customers say  
from what they do  
in the real world**

# What's a rapid experiment?



timeboxed



off the shelf parts  
(little or no programming)



experience,  
not feedback

# Use rapid experiments to...

## Evaluate an opportunity:

- test if people's behavior matches your expectations
- understand what features should be included
- check if the idea is going in the right direction
- understand how people will engage with your idea in the absence of a working prototype
- flesh out how this feature / idea should work to create a strong PRD
- learn something totally unexpected and new



# How to create a rapid experiment

# The Rapid Experimentation Process

1. Pick an idea
2. Make a list of all the questions you have about the efficacy of the idea
3. Select the most critical question(s) to focus on
4. Design an experiment to answer your question
5. Create the prototype to support the experiment

## *1. Pick an idea*

**To run an experiment, you must select an actual idea that everyone agrees on.** Don't start to design your rapid experiment until you've determined what the basic idea actually is.



*Tip: Don't move on to Step 2 unless you have an artifact that describes the idea / feature you want to test*

*Artifacts can be storyboards, sketches, a bullet list of features, or even a very rough PRD*

# Example idea: Auto album

New feature for photography library that makes albums from your photos using AI

## Example artifacts:

### Automatic photo album feature

- AI reviews your photos
- Automatically generates photo albums for events, vacations, etc.
- Creates a list of attendees at the event and asks if you want to share the album with those people via email or text

A bulleted list of features or even a rough PRD



A rough sketch or storyboard

# The Rapid Experimentation Process

1. ~~Pick an idea~~
2. *Make a list of all the questions you have about the efficacy of the idea*
3. *Select the most critical question(s) to focus on*
4. *Design an experiment to answer your question*
5. *Create the prototype to support the experiment*

## *2. Make a list of questions*

Ideas we are considering have assumptions about how people will respond, what people will do and how the whole thing will work.

→ **What are the crucial questions that make or break your idea?**

# Example: List of questions for Auto-Album

- Will people want to interact with albums that are auto-generated for them?
- How often will people return to those albums?
- Will people share the albums with other people in the photos?
- Are the types of albums creating the right groupings?
- How editable do people want the albums? Are they a starting point that people can then curate further?
- Is the AI good enough to make effective albums?
- How many albums should be created? What is the cadence?
- Etc...

# The Rapid Experimentation Process

1. ~~Pick an idea~~
2. ~~Make a list of all the questions you have about the efficacy of the idea~~
3. Select the most critical question(s) to focus on
4. Design an experiment to answer your question
5. Create the prototype to support the experiment

### *3. Select the question to focus on*

- Select the **question that is most critical to answer first about the efficacy of this idea/feature for users**
- Focus on validating a match between the user need and solution
- Focus on a key, specific aspect of the idea and a specific behavior
- NOT...
  - too broad and high level “Does everyone like it?”
  - too technical “Is it technically feasible?”
  - too logistical “How will I implement it exactly?”
  - too much about known facts “Do people take photos?”

# Example: Select a question (or two)

*Specific behavior*

- **Will people want to interact with albums that are auto-generated for them?**
  - How often will people return to those albums?
  - Will people share the albums with other people in the photos?
- Are the types of albums created in the right groupings?
- How editable do people want the albums? Are they a starting point that people can then curate further?
- Is the AI good enough to make effective albums?
- How many albums should be created? What is the cadence?
- Etc...

*One part of  
the idea*

# The Rapid Experimentation Process

1. ~~Pick an idea~~
2. ~~Make a list of all the questions you have about the efficacy of the idea~~
3. ~~Select the most critical question(s) to focus on~~
4. *Design an experiment to answer your question*
5. *Create the prototype to support the experiment*

## *4. Design an experiment to answer your question*

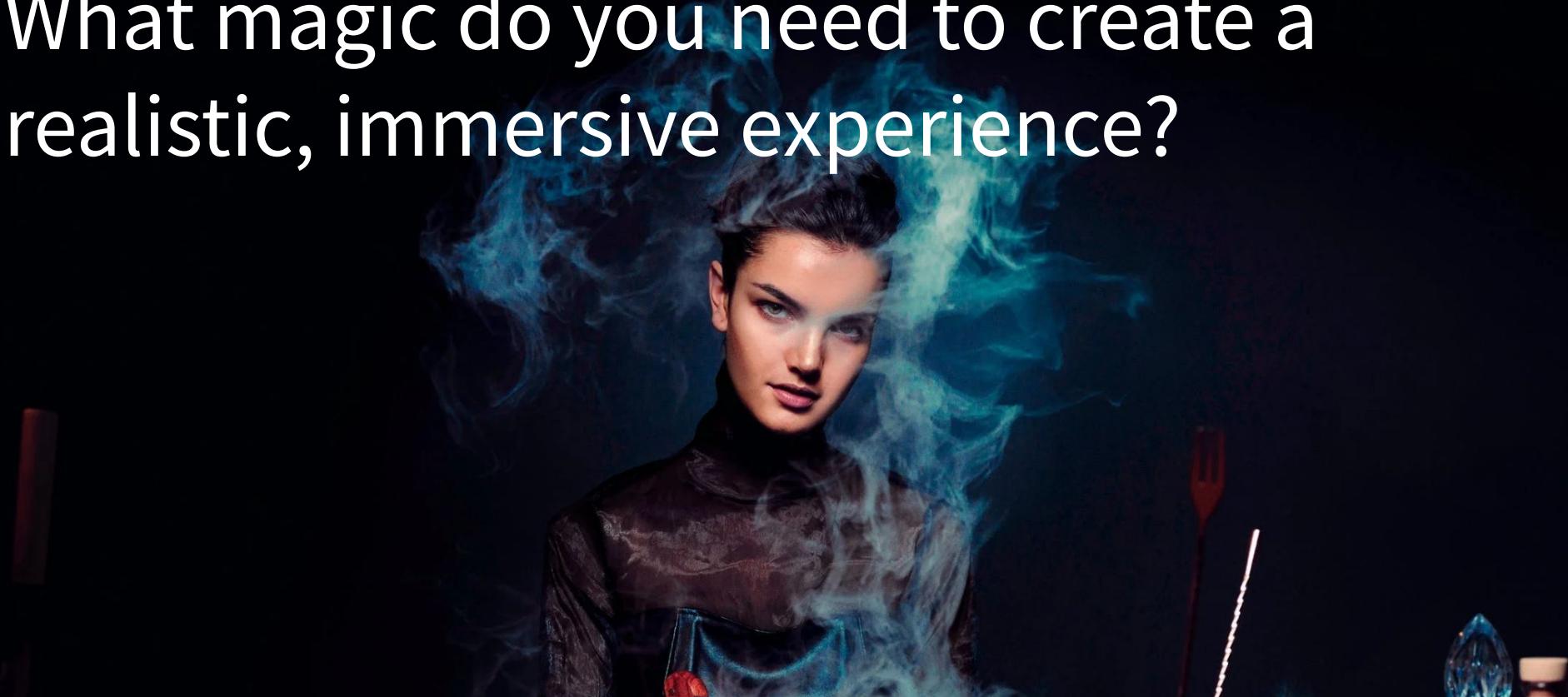
**Goal:** evoke “real” behaviors in “real” situations

- the scenario must be
  - believable
  - immersive
  - natural
- the scenario must allow you to test/measure/observe what you need to (focus on what you want to learn)
- feel free to use “confederates”



# Wizard of Oz prototypes:

## What magic do you need to create a realistic, immersive experience?



# Example: Personalized Dressing Room



**Idea:** Make an appointment to get a dressing room full of clothes put together by an AI personal shopper.

**Question:** Will women in their 20s enjoy an alternative personalized shopping experience?

**Experiment:** Recruit 10 target customers. Have participants fill out an online questionnaire via Google Forms that “informs an AI”. Ask them to include photos of 3 favorite outfits. Fill dressing room of clothes for them in store at an appointment time.

# Example: Chore management system

Happy Thursday! Your chore is:  
Sweep Stairwells 12-13 and 23-24.

When your chore is complete, please  
reply with: Y (Yes, it's done) or N (No,  
I can't do it today).

You can also send a comment.

Today 8:00 PM

Have you completed your chore  
today?  
You can reply with: Y (Yes, it's done)  
or N (No, I can't do it today).

Yes

Great, thank you!

**Idea:** System for residents at a group home to manage daily chores and other community activities.

**Question:** Will residents of group homes complete more chores in a timely manner if they receive reminders?

**Experiment:** Recruit 9 residents at 3 different group homes. Send out text reminders via Google Voice by hand in the morning. Send out additional reminder if no response in the evening. Forward completion information to group home coordinator when it is received.

# Example: Different size bodies



**Idea:** Show different sizes of jeans on a product page (not just small sizes) to help women of all sizes feel included on a retail shopping site.

**Question:** Will seeing jeans on different sizes of women lead to an increase in sales?

**Experiment:** Recruit employees in a variety of sizes who would feel comfortable being photographed. Put up a white sheet and have each employee pose for a quick photo in a pair of jeans (select the most popular type) in their size. Create an image of all the photos and post it to the public online shopping site. Measure impact on sales over the course of a week.

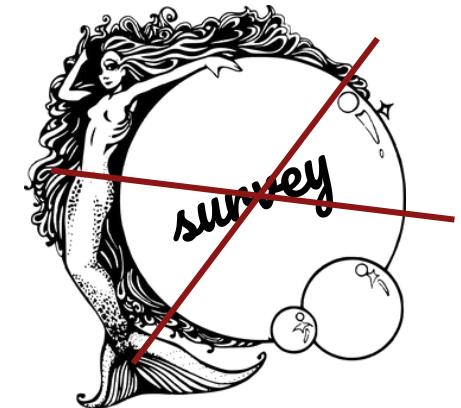
# Skits are great VERY rapid experiments



**survey ≠ experiment**

people are notoriously bad at  
predicting (and remembering)  
their own behaviors

**beware their siren call**



# Example: Design an experiment: Auto Folders

**Idea:** Auto generate photo albums for people via AI

**Question: Will people want to interact with albums auto-generated for them?**

How often will people return to those albums?

Will people share the albums with other people in the photos?

**Experiment:**

- Recruit 5 people and ask them to share their Google Photos with the research team.
- “Wizard” will go in on a daily basis and create folder(s) from photos.
- Will also send text messages to the participants telling them about folders that were created and asking them if they want those photos shared with a list of people in them.

# The Rapid Experimentation Process

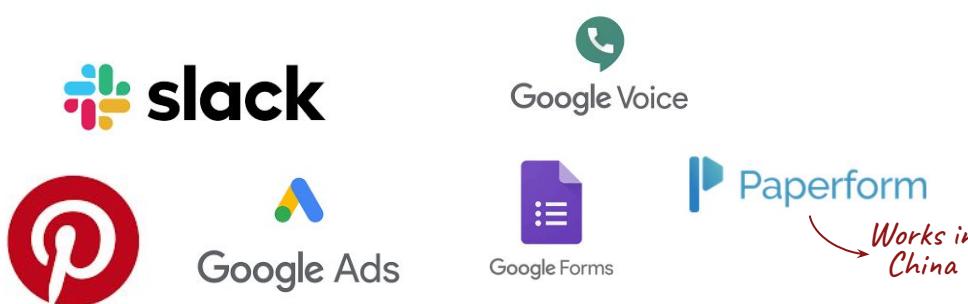
1. ~~Pick an idea~~
2. ~~Make a list of all the questions you have about the efficacy of the idea~~
3. ~~Select the most critical question(s) to focus on~~
4. ~~Design an experiment to answer your question~~
5. ~~Create the prototype to support the experiment~~

## *5. Create the prototype to support the experiment*

- Create the prototype and plan for...
  - measurement
  - logistics
  - follow-up interviews
  - sharing the data
  - ethical considerations

# Create the prototype

- What can you use that's off the shelf?
- Is there anything you can automate without coding?
- What should the instructions be?
- How can you be quick and scrappy?
- Example [YouTube video](#) with Wizard of Oz Prototype



# Example: Create the prototype: Auto-folders

- Google sheet with list of folder types we are creating
  - By location
  - By date/time of event
- Use Google Voice text messaging to send SMS messages to people about sharing albums automatically
- Use Google Sheets to track who is going to be the wizard each day and track progress
- Experiment will run for 1 week

# Plan for measurement

People are notoriously bad at predicting (and also remembering/reporting) their own behaviors

- **What can you actually observe yourself?**
- **What can you actually measure?**
- **How will you follow up?**



# Example: Plan for measurement: Auto-folders

1. **Track if participants click to view folders** from the SMS links they are sent and if behavior tapers off over time
2. Have participants keep a **journal of their interaction** with the folders -- have them note in the digital journal (a google form) when they interacted with an automatically generated folder
3. **Track if participants want folders shared** with people that they are asked about
4. Create a guide with questions for a **follow-up interview** with each person

# Plan for logistics

- **Run a pilot first. Always run a pilot. Don't ever proceed without a pilot.**
- Create a back-up plan
- Example: Auto-folders back-up plan
  - Identify backup person in case a wizard becomes unavailable
  - Offer email as a communication method if SMS won't work for a participant

# Plan for follow-up interviews

- **Learn about the experiment**
  - Start very open-ended
    - “How did that go?”
    - “What worked?”
    - “What didn’t work?”
  - Ask for specific stories: “What’s an example of X?”
  - Ask about any specific areas of interest at the end:  
“What did you think about when you got a message  
that said XYZ?”

# Plan for follow-up interviews

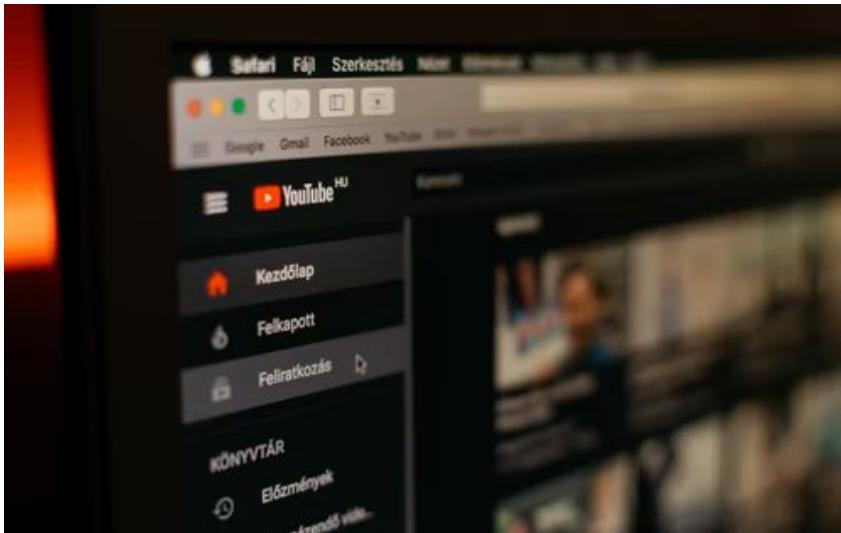
- **Learn about the problem**
  - “How does this compare to other things you use?”
  - “When did you use this? What else were you doing at that time?”
  - “Tell me about your family / work / hobby...”
    - anything relevant to understanding their context



*Tip: Create an interview guide with questions that you will ask in person (not in a survey or in an email!)*

# Plan for sharing the data

“This is quote from a subject....”

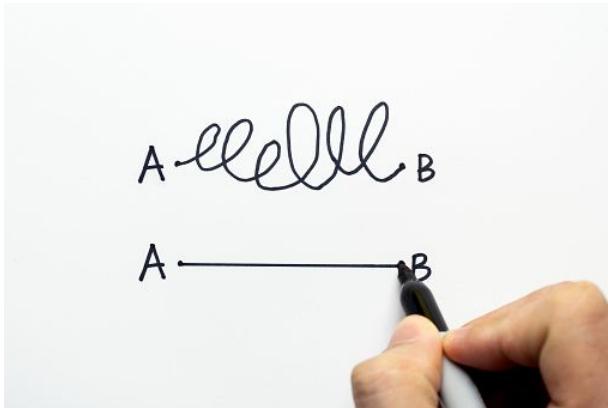


# Plan for ethical considerations

- Deception
- Privacy
- Bystanders
- Awareness of harm



# 3 KEY tips



When in doubt,  
keep it simple



Plan for logistics



Prepare to be surprised

# The Rapid Experimentation Process

1. Pick an idea
2. Make a list of all the questions you have about the efficacy of the idea
3. Select the most critical question(s) to focus on
4. Design an experiment to answer your question
5. Create the prototype to support the experiment

# Experiment Planning Worksheet

## EXPERIMENT PLANNING WORKSHEET

### BIG IDEA

### QUESTION

### HYPOTHESIS

### EXPERIMENT OVERVIEW

[The 2-3 bullet description of what is going on goes here.]

### DETAILED EXPERIMENT DESIGN

1. **Participants** (description of target participants, how many, recruiting strategy, subject compensation/value plan, link to screener)

2. **Preparing for study & prototype creation** (including any required software, supplies, additional helpers, prototypes, design work, links to any related documents, etc.)

3. **Running study** (length of study, expected plan for management during study, backup plan)

4. **Analysis plan** (including link to an interview guide for the follow-up interview)

5. **Artifact collection plan** (how will you document what happened? Audio? Video? Photos?)

### OPEN ISSUES

# EXPERIMENT PLANNING WORKSHEET

**BIG IDEA**

**QUESTION**

**HYPOTHESIS**

**EXPERIMENT OVERVIEW**

**DETAILED EXPERIMENT DESIGN**

1. Participants (target participants, how many, recruiting strategy, compensation,

## **EXPERIMENT OVERVIEW**

[The 2-3 bullet description of what is going on goes here.]

## **DETAILED EXPERIMENT DESIGN**

- 1. Participants** (description of target participants, how many, recruiting strategy, subject compensation/value plan, link to screener)
  
- 2. Preparing for study & prototype creation** (including any required software, supplies, additional helpers, prototypes, design work, links to any related documents, etc.)
  
- 3. Running study** (length of study, expected plan for management during study, backup plan)
  
- 4. Analysis plan** (including link to an interview guide for the follow-up interview)
  
- 5. Artifact collection plan** (how will you document what happened? Audio? Video? Photos?)

## **OPEN ISSUES**

# EXPERIMENT PLANNING WORKSHEET

## BIG IDEA

AI creates automatic folders for your photos and offers to share the photos with friends and family featured in the media.



## QUESTION

- Will people want to interact with albums that are auto-generated for them?
  - How often will people return to those albums?
  - Will people share the albums with other people in the photos?

## HYPOTHESIS

Customers will click on albums that are auto-generated to review the content at least once after they are created. They will want to share albums from group events with other people.

## EXPERIMENT OVERVIEW

Recruit 5 people and ask them to share their Google Photos with the research team. "Wizard" will go in on a daily basis and create folder(s) from photos. Wizard will also send text messages to the participants telling them about folders that were created and asking them if they want those photos shared with the people in them.

## DETAILED EXPERIMENT DESIGN

**1. Participants** (description of target participants, how many, recruiting strategy, subject compensation/value plan, link to screener)

- 5 people
- Target participants
  - Mix of ages, genders, geography
  - Regularly take photos on their phone
  - Attend social events where they take photos
  - One extreme user: Professional photographer
- Recruiting strategy
  - Online panel: Respondent.io
  - Friends and family snowball recruit
- Subject compensation
  - \$150 for one week of participation and follow up interview
- Screener

**2. Preparing for study & prototype creation** (including any required software, supplies, additional helpers, prototypes, design work, links to any related documents, etc.)

- Create Google sheet with list of folder types we are creating
  - By location
  - By date/time of event
- Manage sharing albums via text message
  - Set up Google Voice text messaging to send SMS messages to people about sharing albums automagically
  - SMS message will ask if users want photos shared with certain people
    - Template for SMS text message - to be designed
      - Type 1: Album created
      - Type 2: Who to share it with
  - If a participant replies yes to sharing, we will need to collect contact information for people to share with via SMS (or get it from Google photos contacts?) and send text/email to the person to be shared with

**3. Running study** (length of study, expected plan for management during study, backup plan)

- Experiment will run for 1 week
- Use Google sheet to track who is going to be the wizard each day and track progress
- If participants are non-responsive or do not upload to Google Photos regularly, give them SMS reminders each day.

**4. Analysis plan** (including link to an interview guide for the follow-up interview)

- Track if participants click to view folders from the SMS links they are sent and if behavior tapers off over time
- Have participants keep a journal of their interaction with the folders -- have them note in the digital journal (a google form) when they interacted with an automatically generated folder
- Track if participants want folders shared with people that they are asked about via SMS
- Create a follow-up interview guide with questions for a follow-up interview with each person
- Schedule follow-up interviews with each subject

**5. Artifact collection plan** (how will you document what happened? Audio? Video? Photos?)

- Take screenshots of text message exchanges
- Video tape follow up interviews

**OPEN ISSUES**

1. How will we get subject contact information? Should subjects share their Google contacts with us?
2. How will folders be named?
3. How long will it take to do the auto folder creation for the wizard?
4. What time of day should SMS messages go out?
5. Participants may be hesitant or uncomfortable to share photos with strangers. How do we ensure we get representative, naturalistic data?
6. Participants might feel pressured to share photos they might not otherwise share with strangers. How do we ensure their privacy concerns are met?
7. How do we anonymize the participant from the wizard e.g: blinding procedures?

G<sub>2</sub> O<sub>1</sub>  
F<sub>4</sub> O<sub>1</sub> R<sub>1</sub>  
I<sub>1</sub> T<sub>1</sub>

# Next up: Recruiting

- **Participants** is a key bullet on the worksheet.
- In the next section, let's talk about how to recruit the right participants...





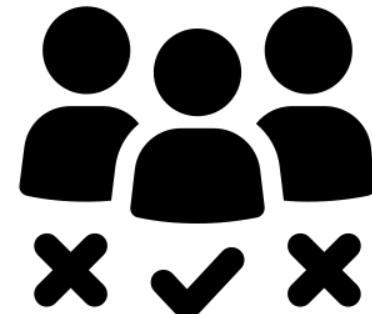
# Subject recruiting

# Your goal

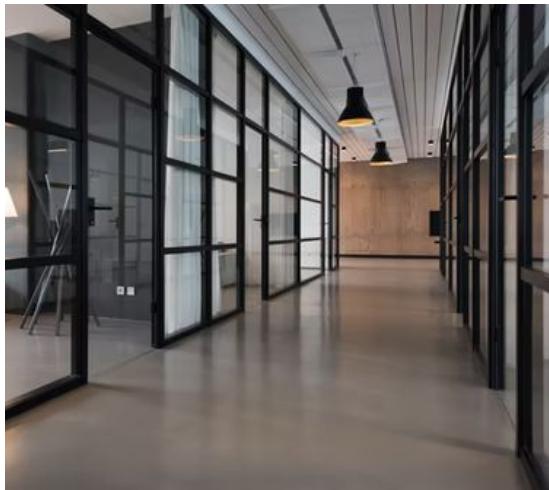
**Screen IN good matches**

**Screen OUT bad matches and bad actors**

- Talk to a semi-representative sample of people you are actually targeting
- Be inspired by a range of behaviors & attitudes
- Avoid false negatives and positives



# Bias & bad actors: A few stories



“The Hallway Bias”



“The Spy”



“The Liar”

# Deliberate screening

**Subject criteria you identify as being critical to study participation**

Can assess a match formally with a questionnaire (screener) or informally through discussion

The image shows a digital form with four questions:

- Is there a dog that lives in the home with you? \***  
Two radio button options: Yes and No.
- How many dogs live in your home? \***  
A short answer text input field.
- What is the age of your youngest dog? \***  
A short answer text input field.
- On average, how many hours per day do you spend in the same location as your dog? \***  
A short answer text input field.

# FIRST: Identify desired / undesired criteria

- Behaviors or attitudes (loves math, never exercises)
- Tool/tech knowledge (tech guru, PPT expert, never used G Drive)
- Domain expertise (e.g. for meditation, avoid yoga instructors)
- Demographic characteristics (gender, age, income)
- Anything you want or don't want subjects to have or do



# Avoid bias in a formal screener

- **Be vague** about the goals of the study - you don't want people to game your screener
- **Don't make the right answer to each question obvious --** obfuscate what you are looking for, don't ask yes/no if it is obvious you are looking for one or the other
- **Don't try to change or influence behavior** from your screener (e.g. ask people in the screener to try your new app)
- **Do not to include your teammates**, yourself, or anyone who knows about the setup and goals of a study

# You try it

*Think and then write down your answer:*

**How would you screen for people who frequently play tennis?**

What kind of question would you ask?

# Example tennis screening question

How frequently do you engage in the following activities?

	Never or rarely	Once/year	Several times/year	At least once/month	At least once/week
Soccer	<input type="radio"/>				
Tennis	<input type="radio"/>				
Visit a gym	<input type="radio"/>				
Hiking	<input type="radio"/>				
Rock climbing	<input type="radio"/>				
Basketball	<input type="radio"/>				
Video games	<input type="radio"/>				
Watching TV	<input type="radio"/>				
Reading books	<input type="radio"/>				
Yoga	<input type="radio"/>				

- Question **gauges tennis frequency without giving away the topic of the study**
- Offers a way to **assess general activity / interests** without asking directly.
- This question was created in Google forms.

# Sample complex screening criteria

Gender	<ul style="list-style-type: none"><li>• Male and female</li></ul>
Employment	<ul style="list-style-type: none"><li>• Must be employed</li><li>• No employees of Acme or its competitors</li><li>• No market researchers</li><li>• No military</li><li>• Not participated in market research in the last 6 months</li></ul>
Age	<ul style="list-style-type: none"><li>• 21-40</li></ul>
Financial	<ul style="list-style-type: none"><li>• \$25,000-\$55,000 annual income</li><li>• Finances are tight / had a problem paying for something large in the past</li><li>• Credit score &lt;580</li><li>• Have used a payday lender (quota =1)</li></ul>
Education	<ul style="list-style-type: none"><li>• A mix of high school and college</li><li>• Screen out higher ed</li></ul>
Technology	<ul style="list-style-type: none"><li>• Has a relatively recent smartphone</li><li>• Exclude people who are extremely tech savvy</li></ul>
Shopping profile	<ul style="list-style-type: none"><li>• Must shop for themselves</li><li>• New to Acme</li></ul>
Interests	<ul style="list-style-type: none"><li>• Would like to purchase: electronics, video games, tv, headphones, iPad</li><li>• Shopping at Macys, Footlocker, JCPenney, Best Buy, Walmart, Target</li></ul>
Online shopping	<ul style="list-style-type: none"><li>• Spend at least \$150 on online shopping in the last 6 months</li><li>• Discount shopper</li><li>• Mix of habits around browsing/purchasing online, computer/tablet v. smartphone</li><li>• Exclude people who never shop (browse) online</li></ul>
Family	<ul style="list-style-type: none"><li>• Married / single (mixed)</li><li>• Has 1+ children living at home? (quota =2)</li></ul>

# A few tips for screeners



## Make it short and mobile friendly

Make the screener as short as possible and mobile friendly (lots of multiple choice and low free responses)



## Compensate subjects or consider the value

Tell subjects what the value will be for them if they participate in your study. Will they be paid? How much? Are they contributing to the greater good? Will they get anything out of it? Are you going to help them solve a problem?



## Survey Monkey Question Bank

Stumped for how phrase a question? Survey Monkey has a searchable question bank.



## Screen for good communication

For studies where you are not recruiting people you know in the dorm, follow up with a phone call after they fill out the screener and ask a few open ended questions to assess their language & communication, likelihood to be an axe murderer, etc.

Alternatively, ask an open-ended question at the end to gauge communication.

# Where to recruit



## Social Media

Facebook, Instagram,  
LinkedIn, Twitter, Nextdoor



## Online affinity groups

Places where people digitally hang out including discussion boards, reddit channels, meetups, Facebook groups, Pinterest boards, Stanford groups



## Online panels

Respondent.io,  
SurveyMonkey panels



## Personal networks

Clubs, alumni lists, lists at work, friends of friends of friends (snowball recruit)

Physically GO to where people who hold the positions you want to engage with and invite them (i.e. coffee shop employees or patrons)

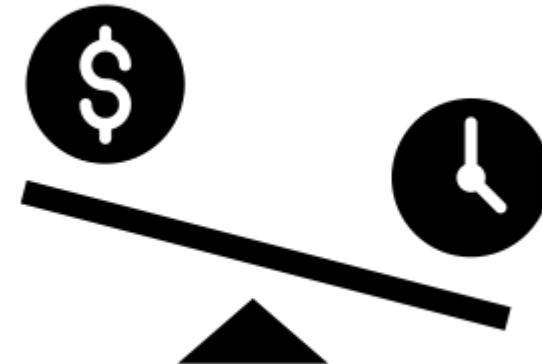
# Subject compensation costs

- Consumers are typically compensated \$125-175 for an hour of their time and goes up from there
- Enterprise participants or very specific scavenger hunt recruits can be \$250-500/hr
- Longitudinal studies range from \$150-1,000 depending on how much time is involved
- Friends can be compensated with chocolate or Starbucks gift cards for quick studies
- Amazon gift cards, Venmo and Paypal work great for remote compensation

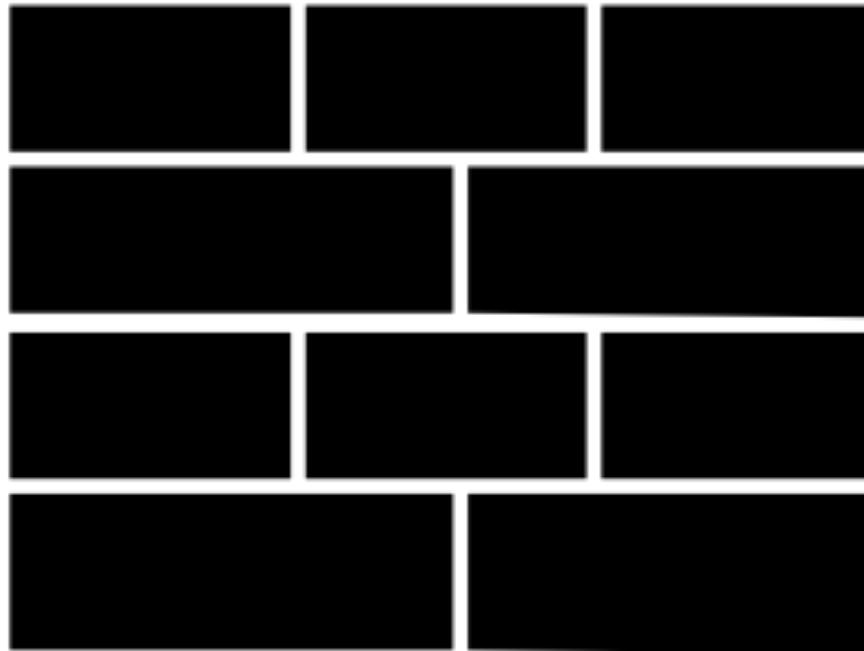


# Subject recruiting & compensation ethics

- Status disparities
- Indirect coercion
- Uneven value



# Brick wall: Is someone better than no one?



# Example: Recruiting for Auto Folders experiment

- 5 people
- Target participants
  - Mix of ages, genders, geography
  - Regularly take photos on their phone
  - Attend social events where they take photos
  - One extreme user: Professional photographer
- Recruiting strategy
  - Online panel: Respondent.io
  - Friends and family snowball recruit
- Subject compensation
  - \$200 for one week of participation and follow up interview
- Screener

# Experiment Planning Worksheet

## EXPERIMENT PLANNING WORKSHEET

### BIG IDEA

### QUESTION

### HYPOTHESIS

### EXPERIMENT OVERVIEW

[The 2-3 bullet description of what is going on goes here.]

### DETAILED EXPERIMENT DESIGN

1. Participants (description of target participants, how many, recruiting strategy, subject compensation/value plan, link to screener)
2. Preparing for study & prototype creation (including any required software, supplies, additional helpers, prototypes, design work, links to any related documents, etc.)
3. Running study (length of study, expected plan for management during study, backup plan)
4. Analysis plan (including link to an interview guide for the follow-up interview)
5. Artifact collection plan (how will you document what happened? Audio? Video? Photos?)

### OPEN ISSUES



# Synthesizing and sharing results of rapid experiments

You recruited your subjects, planned and ran the study. Now what?

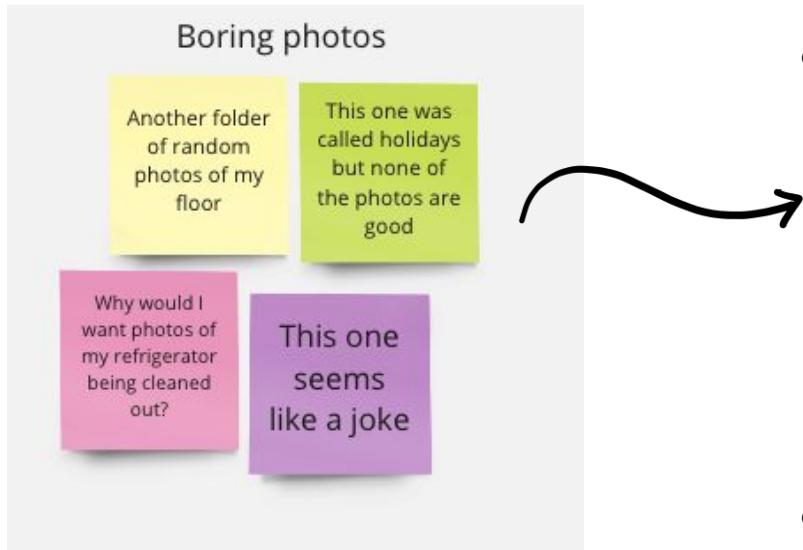
# Synthesize the data

- **Gather:** All the data and post-it notes or an online white boarding app (Miro / Mural)
- **Write:** Note observations and learnings from the data on the sticky notes (1 observation / learning per note)
- **Group:** Start grouping the stickies into affinity groups and name each grouping



# 2 steps to insight...

## What did we observe?



## What might all this mean?

- What are the **top interpretations** / concerns?
  - What have we **learned about this idea?**
  - What have we **learned about the problem?**
    - What new insights have emerged about the problem you are solving and the needs of the people involved?
    - What is actually important to people?
- **Where should we go from now?**

Review and cluster observations

# Sharing the data

- Take stakeholders on the **journey of what happened** – not just a data dump
- Make a **one pager** that...
  - Explains the experiment (briefly)
  - Share key insights
  - Include quotes, images, the prototype
  - Introduce next steps



# One pager: Auto folders

**Big Idea:** Auto-generation and sharing of photo albums for your photos.

**Question:** Will people want to interact with albums that are auto-generated for them?

**Experiment description:** Recruit 5 people and ask them to share their Google Photos with the research team. “Wizard” will go in on a daily basis and create folder(s) from photos. Wizard will also send text messages to the participants telling them about folders that were created and asking them if they want those photos shared with the people in them.

## Learnings

1. People were **disappointed in the folders** that were created as many of the photos were not that interesting or not personalized enough
2. **Few people wanted to share** auto generated folders with friends.
3. Users often had **multiple photos for one event that were repetitive**.  
The real benefit would be helping them select the best one -- not just grouping them together.

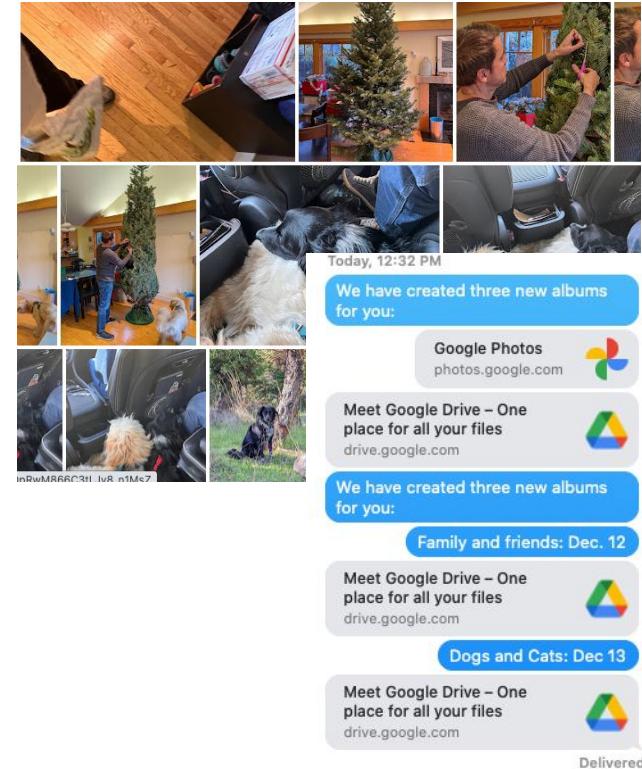
## Next steps

1. Evaluate if it's possible for AI to help select the “best” photo from a set of options
2. Consider experimenting with photo selection options in lieu of auto folders

## Quotes:

*“This is an album of random stuff from Thursday. Half the stuff I would just throw out.”*

*“I have so many photos that more photo albums were not as useful as I thought ...  
What I really want is help choosing the best photo to post on Insta from whenever I go out!”*



# Chore management system

**Big Idea:** System for residents at a group home to manage chores and community activities.

**Question:** Will residents of group homes complete more chores in a timely manner if they receive daily reminders that they must respond to?

**Experiment description:** Recruit 9 residents at 3 different group homes. Send out text reminders via Google Voice in the morning. Send out additional reminder if no response in the evening. Forward completion information to group home coordinator when it is received.

## Learnings

1. More residents completed their chores in a timely manner when they had daily reminders. Success!
2. Residents loved getting reminders so that they didn't have to look up their assignment. Numerous participants reported that they finished their chore sooner than usual and it was a load off.
3. Some reported that it helped their relationship with their roommate since a neutral third party was doing the nagging.

## Next steps

1. Create a pilot that is automated with all residents in a home.
2. Run an experiment with reminders beyond chores. Is multiple reminders too much?

## Quotes:

*"It is a friendly reminder and provides weekend contact with staff offsite that I miss."*

*"Got to do my chore and finish it sooner than usual"*

*"I didn't have to fight with my wife about it this week! It just happened."*

Happy Thursday! Your chore is:  
Sweep Stairwells 12-13 and 23-24.

When your chore is complete, please reply with: Y (Yes, it's done) or N (No, I can't do it today).

You can also send a comment.

Today 8:00 PM

Have you completed your chore today?  
You can reply with: Y (Yes, it's done) or N (No, I can't do it today).

Yes

Great, thank you!

# Key Takeaways

1. Rapid experimentation is a user research method that helps product managers **understand how people might actually engage with a new idea** in the real world before it actually exists
2. Planning and running a rapid experiment is a **five step process**:
  - Pick an idea
  - Make a list of all the questions you have about the efficacy of the idea
  - Select the most critical question(s) to focus on
  - Design an experiment to answer your question
  - Create the prototype to support the experiment
3. The **Rapid Experiment Worksheet** is a great resource for planning and managing experiments — especially for managing all the logistics
4. **Subject recruiting** is a deliberate process for gathering a representative sample of participants for your study while screening IN good matches and screening OUT bad matches and bad actors
5. When sharing the results of a Rapid Experiment, **take stakeholders on the journey** of gathering data in addition to sharing the final learnings

# **Part 3: Deep Dive moderated usability testing**

Stanford Center for Professional  
Development

Testing the **usability** of a design idea by showing people **prototypes**, asking them to imagine realistic **scenarios** and complete specific **tasks**



# Use moderated user tests to...

**Get feedback on specific options for designing the experience**

- understand if the design is going in the right direction
- identify usability and workflow issues as soon as possible
- ensure that the final design is easy to use
- compare multiple design ideas and select the best direction
- refine a design concept

# Moderated user tests vs. rapid experiments

	<b>Moderated User Test</b>	<b>Rapid Experiment</b>
<b>Moderator's Role</b>	moderator directly interacts with user to facilitate experience	moderator sets up experience but does not really interact with the subject as a moderator (can be a confederate)
<b>Scenario Reality</b>	subject is asked to imagine scenario is real	subject is placed in realistic scenario (and may fully believe it is real)
<b>Measurement Flexibility</b>	in addition to planned measurements, experimenter can probe to understand behavior real-time	outcome is typically measured by planned, objective measurements with a follow-up interview to probe what happened after the fact
<b>Data Signal</b>	what people think they might do	what people actually do
<b>Feedback Granularity</b>	detailed feedback on small nuances of an interaction and workflow steps	big picture feedback on the overall experience, context of use, personal impact, and market fit of a new idea

# Questions each approach answers

## Moderated User Test

### Best for Idea Refinement

- Is this specific design working for people from a usability perspective?
- Why isn't anybody interested in this solution?
- What is missing from this design?
- Why can't people complete the task?
- How do people feel about this interaction?
- What are people thinking about when they are making choices between A & B?
- What *might* people do with this?
- How easy is this to use?

## Rapid Experiment

### Best for Concept Testing

- What would people do on their own?
- Is this idea interesting to people? Will they engage?
- Are people actually using this to solve their problems successfully? Is this addressing a need?
- Do I have the minimum set of features so that people can complete a task on their own?
- Given the option between A and B in the wild, which one would someone choose?

# Experiment to reality continuum

**100% real scenario  
and product/service**

**Imagined scenario with  
rough paper prototype**



Totally experimental

Totally moderated

## What questions are you trying to answer?

- What stage of product development are you in?
- What is possible to prototype?
- How important is it to be real?
- What's the right sensor?

# Elements of a moderated test

A few things need to be in place to run a moderated test:

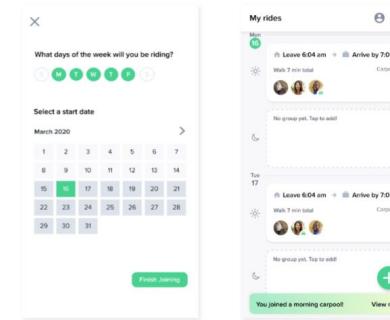
- **Prototype:** You need something to test. This could be a paper sketch, a digital prototype of some screens, a physical prototype of product.
- **Moderator:** Someone who is running the study and engaging with the participants.
- **Participants:** People you have recruited to give you feedback. These tests are qualitative, so 3-8 people is about right.
- **Location:** This is either a physical location where you meet up with subjects or a virtual teleconference space for testing.
- **Test plan:** A written plan for the scenarios and tasks for testing (plus follow-up questions).

*Let's look at each element in detail...*

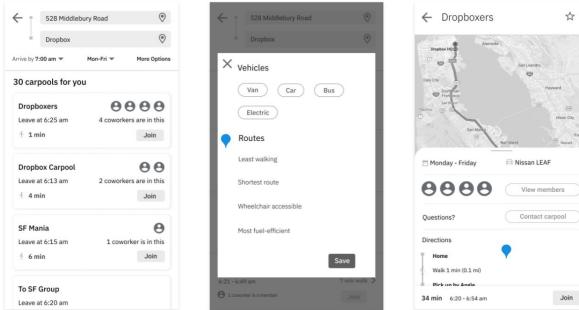
# Prototypes for testing can be varied



Paper Prototype



Visual Design



Grayscale Wireframe



Physical Prototype

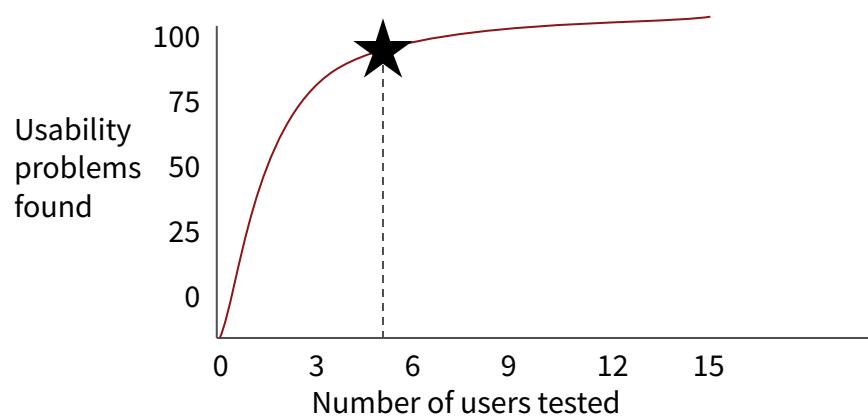
# Moderator

	Professional Researcher	You	A Colleague
Pros	<ul style="list-style-type: none"><li>• Teams using better more professional methodology have been proven to find more usability issues</li><li>• Will ask the right questions phrased in a way to elicit reliable, deep feedback</li><li>• Will pull maximum information out of the study when they synthesize the data</li></ul>	<ul style="list-style-type: none"><li>• Knows the product and the top priorities for feedback</li><li>• Available whenever you want (as long as it's a high priority)</li><li>• First-hand interaction with customers is important for successful product managers</li><li>• <b>Bad user testing beats no user testing!</b></li></ul>	<ul style="list-style-type: none"><li>• Knows the product and the top priorities for feedback</li><li>• May be a subordinate who you can assign tasks to on a timeline that works for the project</li><li>• May have some user research skills that you don't have (never hurts to ask around!)</li><li>• May be a more neutral moderator if they are not personally involved in product design</li></ul>
Cons	<ul style="list-style-type: none"><li>• Can be expensive</li><li>• May not be available</li><li>• May take time to bring up to speed on a complex product / service</li><li>• May be difficult to find someone reliable and highly competent</li></ul>	<ul style="list-style-type: none"><li>• Can introduce bias into the study by asking poor, leading questions</li><li>• May not ask strong questions, missing a chance to uncover key information from participants</li><li>• May struggle to create an appropriate study plan and protocol</li></ul>	<ul style="list-style-type: none"><li>• Same issues as you may have if inexperienced</li></ul>

# Participants

**A small number of participants can find a LARGE number of problems**

According to numerous studies by the Nielsen-Norman group and many others over the last 20 years, the optimal approach is to **test numerous times with 5 users each time**



- **First 5 users will find majority of problems**
- Enough users to see a pattern
- Other problems will be masked by the most egregious ones that must be fixed first to discover more
- No need to keep observing the same problems over and over

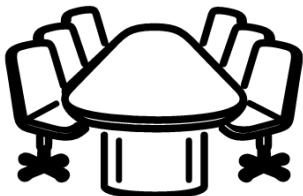
*Instead, test iteratively!*

Adapted From:

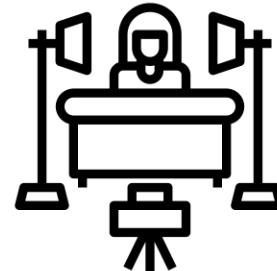
Nielson, J. (2000, March 18). [Why you only need to test with 5 users.](#) Nielsen Norman Group.

# Location options

Conference room



Usability lab



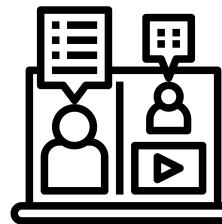
Coffee shop



In home



Online



# Zoom: It's not for just meetings

- A great tool for moderated usability testing remotely
  - Desktop and mobile (Even concurrently!)
  - Recording
  - Mobile mouse control
  - Observers with video off
- Other tools
  - UserZoom
  - Lookback.io



# Tips for success

- **Tech check-in:** Schedule a 15 minute tech check-in the day before the meeting for high stakes tests
- **Reminder emails:** Send out emails the day before the test that remind participants to:
  - Have their laptop and/or mobile device charged or connected to power
  - Have a headset to minimize feedback
  - Make sure they have a WiFi connection
  - Be in a quiet place alone
  - **Not be driving or in a car with other people**

# Administering remote consent forms

- Send a consent form out in advance and ask them to send back a signed copy
- If you haven't already obtained a signed consent form in advance, ask participants to read your consent form out loud, and then verbally give their consent by saying their full name out loud.
- **Pilot signing a consent form** – there are often snafus here!

# Test moderation

The moderation is more  
important than the actual  
questions...so let's start there.

Let's do two quick usability  
tests & compare them

# Test A

# Test B

# What did you notice?

# Top moderation tips

- Observe quietly, stay neutral
- Ask open ended questions and lots of **WHY**
- Encourage thinking out loud
- Don't lead the user or ask leading questions
- Try to understand what was expected vs. what actually happened
- Always blame the prototype, never the user
- Ask about this prototype, not some other hypothetical situation
- Ask the user to think for their situation, not some mythical others
- Act like you have no idea how this works...you want to know how THEY think it should work



# What about planning?

- Introduce what you expect the user to do
- Do not introduce yourself as “The Designer”
- Plan realistic tasks for realistic scenarios
- Have a plan for what you are testing...but veer away from it if the subject is engrossed in the task at hand

**Let's dig deeper in the next module...**

# Test design and recruiting

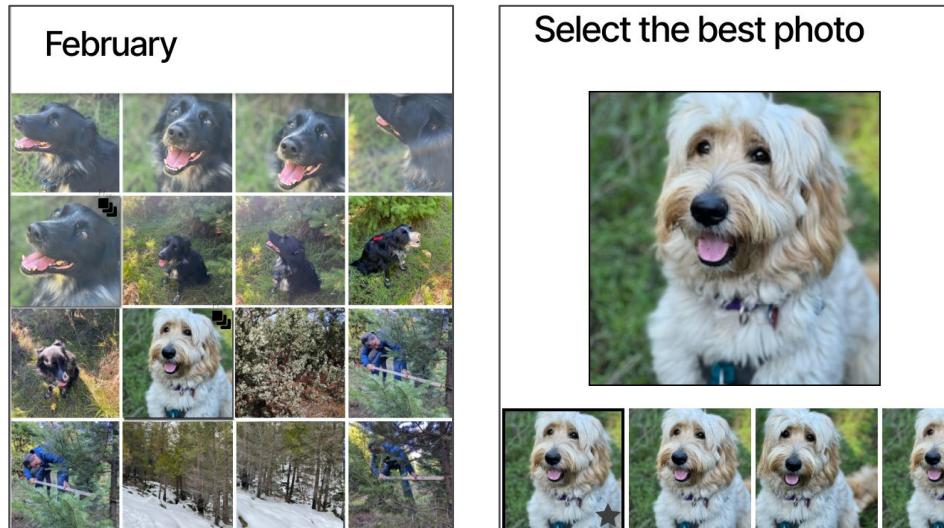
# Our goal with the test design

- Introduce what users might actually do and why
- Create realistic scenarios with realistic tasks
- Balance direction with exploration
- **Find as many problems as possible – this is not a sales effort**



# Let's make a sample test plan for our new Auto Photo Selection feature design

Example design prototype: Photo selection feature



# Step 1: Goal setting

- What is your prototype focused on?
- What do you want to learn?
  - Usability?
  - Workflow?
  - Feature evaluation?
  - Compare two or more ideas?
- What are the top questions to answer?



# Example: Goals for Auto Photo Selection moderated test

- Goal: Usability of key design elements
  - Do users understand that this photo is part of a set?
  - Do users understand how a photo was selected?
  - Can people use the design successfully to select a different primary photo from the batch?
  - Can users select multiple photos?
  - Can users separate the batch?

# Step 2: Target participants

- Who do you want to test with?
  - *Think about the recruiting module from earlier*
- How many subjects? 3-8
  - 3 is absolutely bare minimum but not great
  - Even with small tests, go for diversity
  - **When in doubt, go with 5**
- Consider compensation
- Where will you recruit them? Consider guerilla testing if consumers.

# Example: Target participants for Auto Photo Selection

- Who do you want to test with?
  - People who regularly attend social events where they take photos with their camera
  - People who often take multiple photos
  - People who post photos to social media
  - Range of ages
- How many subjects? 5
- Compensation: \$100 for 30 minutes
- Recruiting plan: Friends and family snowball recruit

# Ethical Considerations



Usability tests can be distressing; **users have left in tears**

If you just include the easiest people to recruit, you are ignoring key customers

Testing/fieldwork can be coercive if there is a power imbalance (e.g., in under resourced communities)

# Ethical Considerations

You have a responsibility to alleviate these issues

- make participation voluntary with informed consent
- let participants know they can stop at any time
- stress that you are testing the system, not them
- make collected data as anonymous as possible

Sometimes must get human subjects approval (IRB)

# Step 3: Tasks and Prompts

- **Are you trying to understand general use OR interaction with specific areas?**
- Plan open observation AND specific tasks
- Create realistic scenarios for people to imagine they are in
- Don't choose tasks that are too fragmented
  - *fragmented = does not represent a complete goal someone would try to accomplish with your application*
- Facilitate **think out loud protocol**
- Keep the test short: 3-4 tasks max, 45-60 minutes

# Start with an open-ended prompt

## Task 1: Explore the screen

Imagine that you just updated a mobile app for storing photos and heard that it had a new feature where it could select the best photo from a set of similar ones. You've finished updating the app, opened it up, and this is what comes up on your screen.

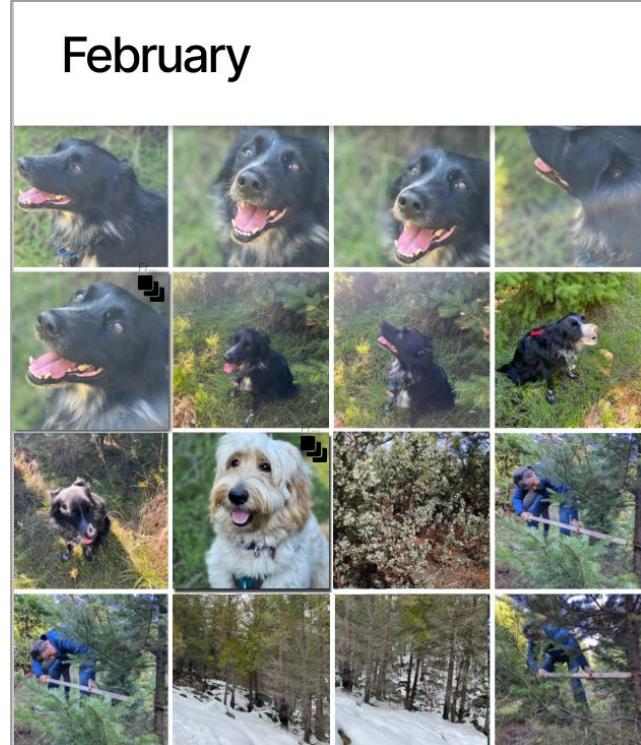
**I want you to take a look and let me know what you think.** And remember to think out loud. *[Pause for exploration]*

What are your first impressions?

# Add specific, realistic tasks

## Task 2: Edit batch selected image

Imagine that you want to choose a different image from the batch that this app put together. Show me how you might do that.



# Experience prototypes

Engineered experiences can be a grey area with rapid experimentation – it's a continuum not a black and white division.



# Step 4: Refine Tasks/Scenarios

- Consider a **natural order** for the tasks
  - At a minimum go from general to specific
  - If users skip the order in your test in the study naturally (and the prototype supports it), let it happen
- **Doing comparisons?** Let people discuss one option, then the second, then compare the two
- For each task, **come up with the questions** you might want to ask to really understand the user's behavior *BUT REMEMBER – AVOID LEADING*

# Example: Refinement/Order

## Task 1: Explore application

Imagine that you just updated a mobile app for storing photos and heard that it had a new feature where it could select the best photo from a set of similar ones. You've finished updating the app, opened it up, and this is what comes up on your screen. I want you to take a look and let me know what you think. And remember to think out loud. [Pause for exploration]

- *What's your first impression?*
- *What do you think is going on here?*
- *What sorts of things do you think you can do in this app?*
- *How might you use it?*
- *What would you do next?*

## Task 2: Edit a batch selected image

Imagine that you want to choose a different image from the batch that this app put together. Show me how you might do that.

- *How did that go?*
- *Anything confusing or frustrating about that?*
- *Is that something that you ever do on your current phone? Why (not)? How does it compare?*
- *What do you think happens to the photo that was selected originally?*

# Step 5: Fill in the details

- Create an intro
- Ask some background questions
  - Understand background of participant and current mindset about this type of product / service / feature
- Add in your tasks/scenarios
- Add wrap-up questions at the end

# Example: Introduction

- We're here to get your feedback on some new ideas for a photo organization app
- We have a prototype of one of the ideas to show you in its early stages to get your initial impressions and will give you some tasks to try out.
- As we go through the task, I want to encourage you to **think out loud**/ If anything is confusing or you don't like it, don't hesitate to let me know.
- I didn't design this product, so you won't hurt my feelings. My only goal today is to get your feedback on what does and doesn't work for you.

# Example: Background Q's

*Let's start with some background questions*

- Tell me a little bit about yourself like where you live, what you do for work, and who is in your family
- What do you use to take digital photos?
- What do you use to store and organize your photos?
- Show me the last photo in your camera roll and tell me about the time that you took it
- How do you organize your photos right now? Show me how you do it on your phone.

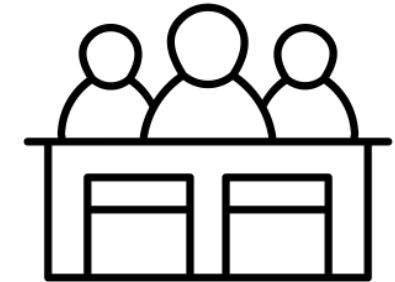
# Example: Wrap-up questions

- How'd that go?
- What do you think about this app?
- Anything especially confusing or that you didn't like?
- Anything that you liked?
- How do you think it compared to what you do currently?
- Thanks for your time!

# A few tips

- Put all of these tasks and scenarios together in one document called a **Test Plan**
- Be prepared to alter scenarios on the fly in response to what happens
- Leave time to update your prototype to accommodate the tasks you're testing
- Ask a lot of why and encourage thinking aloud
- **Run a pilot on a teammate or friend**

# Finally, prepare your observers



**It's great to invite people to observe the study!**

- How many observers
  - In person in a conference room room, 2-3 is acceptable
  - Behind one way glass or virtually (on Zoom) unlimited people can observe!
- Share rules for observing, for example:
  - Join five minutes early
  - Only ask questions when I specifically ask if anyone has any
  - Don't jump in or correct the subject or the moderator
- Provide observers with copy of the tasks that users will attempt and specific issues you want them to watch out for.
- *Consider recording video for people to watch later.*

**That's all you need for a test  
plan!**

# Synthesizing and sharing results

# Synthesizing results

- Just like rapid experimentation, take a look at your data and create affinity groups
- Track what problems participants experienced and why
- Summarize results for each task and for each element of your prototypes

But there's something very important to watch out for...

# What happened here?



Amazon Essentials  
[Visit the Store](#)

Amazon Essentials Women's 2-Pack V-Neck Cl...  
Sleeve Tee Shirt

★★★★★ 24,406 ratings | 119 answered questions

Price: \$8.69 - \$22.08 & Free Returns on some sizes and colors

Fit: True to size. Order usual size.

Size:

Select ▾

Color: Green/Light Pink

Select Size from the left  
to add to Shopping Cart



Share

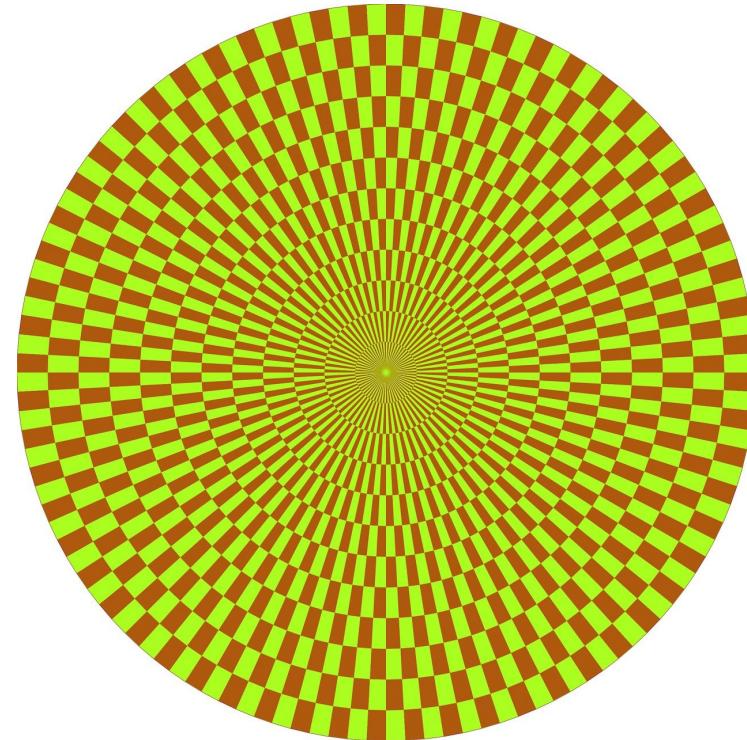


# What usability tells us

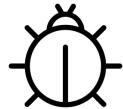
- Usability tests find usability problems - brought to you by the Tautology School of Tautologies
- They do not find deeper needs... unless you look under the covers
- **Is the breakdown we observe a surface-level issue with a quick fix, or indicative of a deeper need?**

# Illusion of progress

Identifying usability issues and fixing them can feel like progress, but could be taking you backwards



# Terminology



**Nit:** surface-level usability issues

vs.

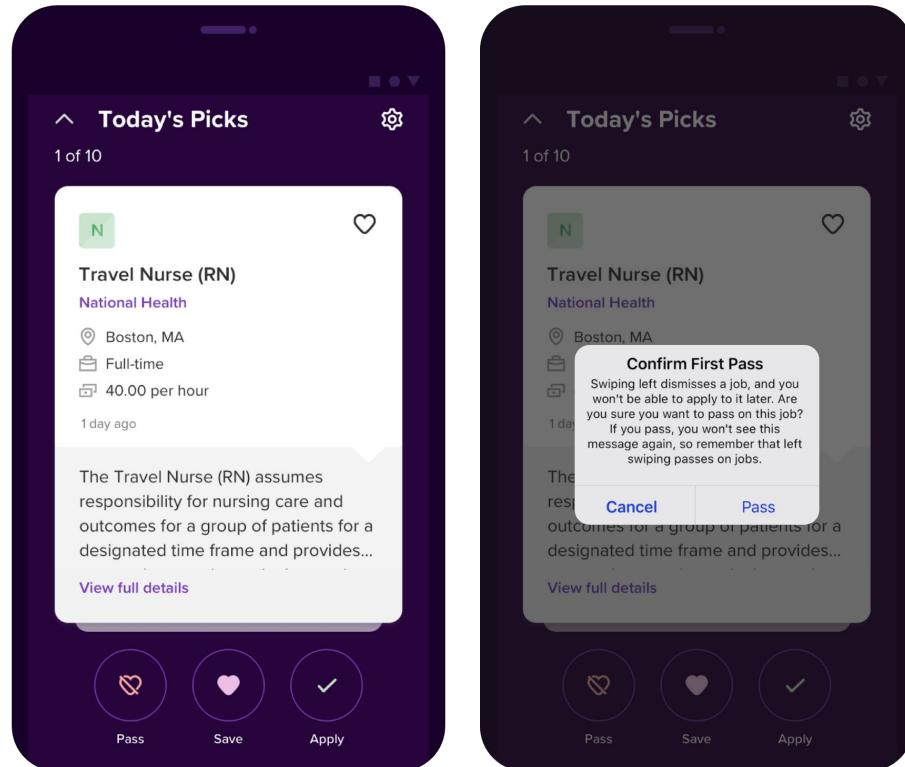


**Need:** surface-level issue is indicative of a deeper mental model misalignment - a reflection of a deeper goal

# Example

Imagine you ran a usability study on this screen in the app Monster (for finding a job)

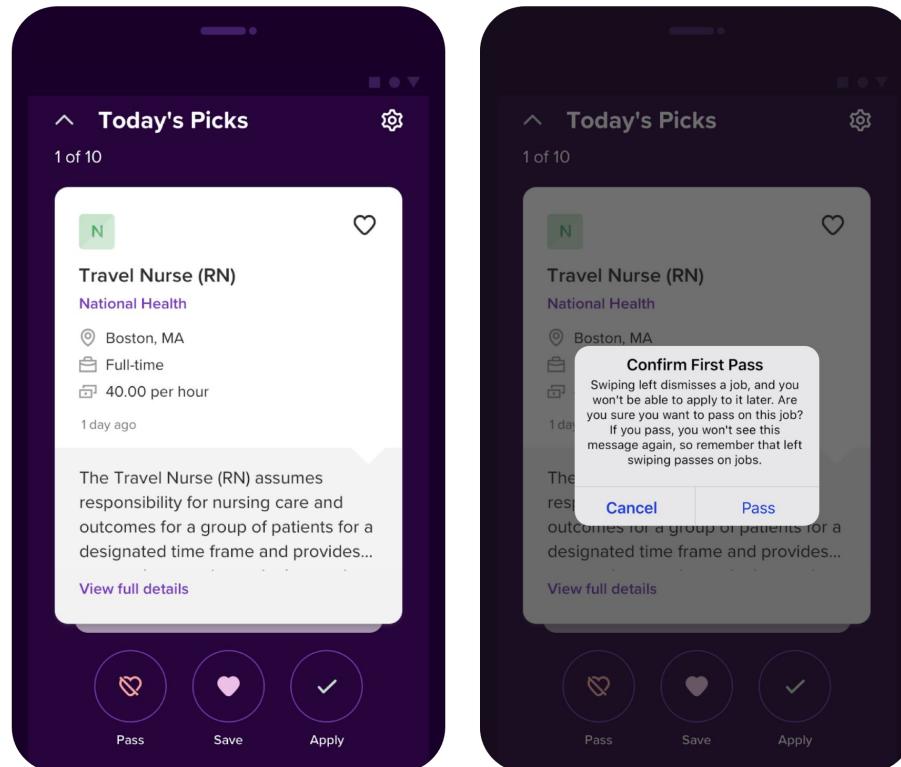
When people click on Pass, they were surprised and unhappy.



# Example

**Nit:** Language is confusing. People are surprised that Pass is going to delete things forever.

**Need:** Users want to consider and save jobs without applying. The Pass feature doesn't address their needs.  
When participants see a job ad, they aren't ready to apply immediately when they are just browsing so they feel stuck. Pass removes the job from their feed forever. Some people in the study took screenshots to save the job. *A larger design issue exists here; different functionality is needed for users to be able to revisit job options.*



# Create a Nits vs. Needs table

Observation	Nit 	Need 	Recommendations / Ideas
The “Pass” button was really confusing for participants. People were surprised at what it did and were unsatisfied with the functionality.	Language confusing. Pass doesn't imply delete forever.	Users want to consider and save jobs without applying. The Pass feature doesn't address their needs.  When participants see a job ad, they aren't ready to apply immediately when they are just browsing so they feel stuck. Pass removes the job from their feed forever. Some people in the study took screenshots to save the job. A <i>larger design issue exists here; different functionality is needed for users to be able to revisit job options.</i>	Consider language "Delete" which is more clear to people  Are cards appropriate for this? Consider having an option for people to Save for later. Reconsider if the overall model of passing makes sense.

# Tips for sharing results

- Prioritize observations
  - **P1: Must fix**
  - P2: Ideally fix
  - P3: Minor annoyance
- Note how many people experienced key problems (observers will always ask)
- Share clearly who you tested on and why
- Include screenshots, quotes and stories to bring your research to life

# Key Takeaways

1. **Moderated User Tests** are a great tool for gathering feedback on design ideas and prototypes where moderators guide participants to try a prototype and answer follow-up questions about it.
2. To run a moderated user test, you need to plan for the following **key elements**:
  - **Prototype**: What you are testing such as a paper sketch, a digital or physical prototype.
  - **Moderator**: Someone who is running the study and engaging with the participants.
  - **Participants**: People you have recruited to give you feedback.
  - **Location**: A physical location where you meet up with subjects or a virtual teleconference
  - **Test plan**: A written plan for the scenarios and tasks for testing (plus follow-up questions).
3. Effective **moderation focuses on eliciting unbiased, deep feedback** from participants by avoiding leading questions and asking a lot of why.
4. Strong **test plans include open ended exploratory prompts** and **specific realistic tasks** ordered in a natural manner.
5. When evaluating study results, consider if an observation is a **surface usability problem** (nit) or a sign of a **deeper issue** in the overall design (need).



# Part 4: Deep Dive Unmoderated Testing

Stanford Center for Professional  
Development

**Method for testing highly  
functional digital products and  
services using online software  
in the participant's natural  
environment without a  
moderator**

# How unmoderated tests work

- **No moderator:** Participants complete the study by themselves with no moderator except for the computer instructions.
- **Discrete tasks:** Participants are given individual predetermined tasks to complete in a certain order.
- **Live sites or robust prototypes:** Unmoderated tests can be carried out on live sites or standalone digital prototypes. They do not work for physical products or paper prototypes.
- **DIY or consultation:** You can plan and conduct your own study or hire the service provider to help you do it for an additional fee.

# Qualitative and/or quantitative output

- **Qualitative output:** Users are recorded as they complete tasks creating videos you can watch.
  - Users can answer predetermined qualitative questions about a task.
- **Quantitative output:** Users' interaction can be measured quantitatively:
  - Task time
  - Task completion
  - Rate of task abandonment
  - Click heatmaps / Clickpaths
  - Rating questions
  - Other types of structured tests: 5-second test, card sort, tree test, etc...



# Integrated tools

- **User pools:** Pre-recruited pools of users you can invite to participate and screen for target demographics, prior experiences, habits or preferences
- **Qualitative analysis:** Tools for tagging, analyzing and sharing recordings
- **Quantitative analysis:** Data visualization tools

# Unmoderated vs moderated tests

	Unmoderated User Test	Moderated User Test
<b>Moderation</b>	no moderator, all questions and plans pre-written	moderator directly interacts with user to facilitate experience
<b>Scenario Reality</b>	subject can be asked to imagine a real scenario although this is often not that effective OR can be captured in the act of actually using a live site and have their real experience recorded	subject is asked to imagine scenario is real
<b>Measurement</b> <b>Flexibility</b>	all questions and measurements are pre-planned; there is no flexibility during a study to adjust plans or ask new questions in response to user behavior	in addition to planned measurements, experimenter can probe to understand behavior real-time and the user's mental model
<b>Data Signal</b>	mix of what people might do and what people actually do depending on the test setup	what people think they might do
<b>Feedback</b> <b>Granularity</b>	detailed, potentially quantitative feedback and tracking of minute aspects of behavior (e.g. mouse click and drag)	detailed feedback on small nuances of an interaction and workflow steps

# Unmoderated test pros and cons

- **Pros**
  - **Fast:** can launch a study and receive results within just a few hours
  - Can run on its own regularly and with multiple people simultaneously
  - Can provide very granular data about behavior
  - Scales to quantitative levels
  - Participants may not feel as self-conscious – you’re not criticizing the designer or someone directly involved with a project
- **Cons**
  - **Limited depth of insight:** no nuance or follow-up on surprising responses / behaviors
  - Can’t help if user gets stuck
  - Early-prototype testing hard without moderator to help participants recover from prototype problems
  - Without a moderator, participants are less engaged in tasks that need creativity, decision making, or careful thought. Participants often race to complete tasks and answer questions briefly because there is no social pressure from a moderator to behave realistically
  - Requires meticulous planning

# Use unmoderated tests to...

- **test a very specific question**
- quickly assess reaction of a large group of participants
- evaluate designs in a completely neutral manner
- track usability evolution over time in a systematic fashion
- accurately assess precise frequency of problems
- convince stakeholders who only believe in quantitative data
- conduct rapid, iterative testing in an agile development environment on the latest build or the live site
- gather feedback on competitors

# How to run an unmoderated test

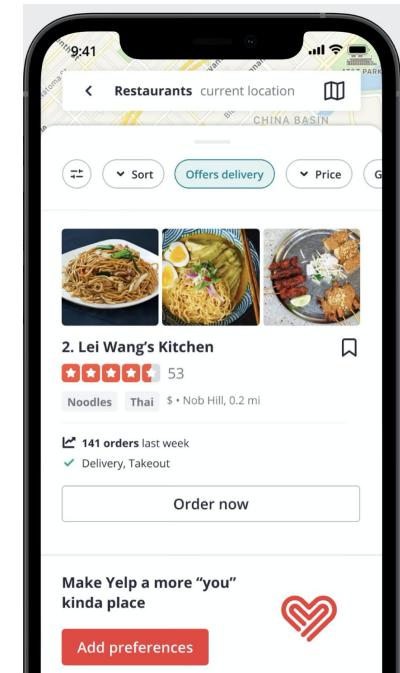
# Elements of an unmoderated test

A few things need to be in place to run an unmoderated test:

- **Prototype / Live site or app:** The artifact being tested must be able to stand on its own two feet without any moderator intervention to make it run.
- **Participants:** Unmoderated service providers usually help recruit consumers to match your criteria (see the module on recruiting). For enterprise users, you usually have to provide the participants.
- **Test plan:** A test plan with scenarios and tasks are key for an unmoderated study.
- **Unmoderated testing software service:** It's important to pick the right provider for your needs as reflected in the test plan.

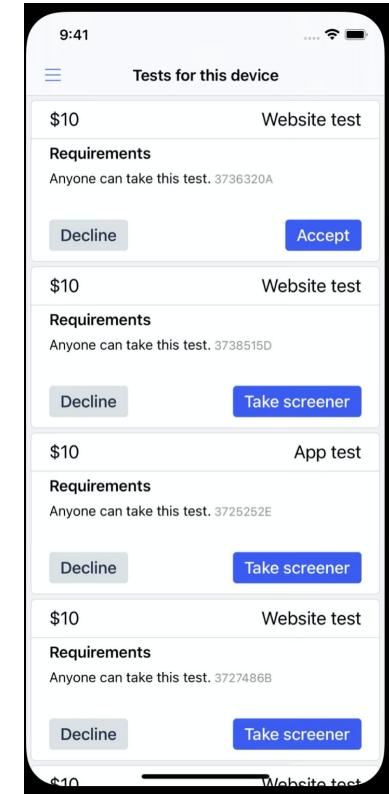
# Prototype or live site/app to test

- Unmoderated tests are often conducted on **live sites or high fidelity prototypes**
- The test can be set up to show **one screen** at a time or a **whole workflow**
- Have an early stage prototype? These are typically tested with moderators so that you can get deeper, more varied feedback on rough functionality and design ideas



# Participants

- Create a **set of recruiting criteria** and a **screener** as described in the subject recruiting module
- Software provider can recruit consumer from their panel / mailing list consumers
  - Most unmoderated software has a screener set up process in their workflow
- Enterprise participants are typically provided by the company running the study
- Quantitative unmoderated studies are usually for consumer products and services due to ease of recruiting



# Creating a test plan for an unmoderated study

# Creating an unmoderated test plan

- Test plans have the same structure and elements as a moderated plan: Intro questions, scenarios and tasks, follow-up questions, wrap-up. **Start with the same process you just learned for moderated testing.**
- Very important to estimate the time for each tasks so that a test can be completed within ~15 minutes (15 min mobile, 20 min desktop)
- Quant or qual test?
  - Quant: Each task should have clear criteria for successful completion.
  - Qual: Think-out-loud protocol is generally for qualitative testing only.
- Instructions must be extremely clear with no nuance; there will be no human helping if someone is confused about what to do

# 1. Create study goals

Study goals will determine what you want tracked and measured in your test and what service to use

**Example goal 1:** Track how long it takes to complete sign-up for your website

## Requirements for software:

- Track time on task
- Display average and individual times for participants
- Export data

# 1. Create study goals

**Example goal 2:** Compare sign-up process design A and B for a native mobile app

## Requirements for software:

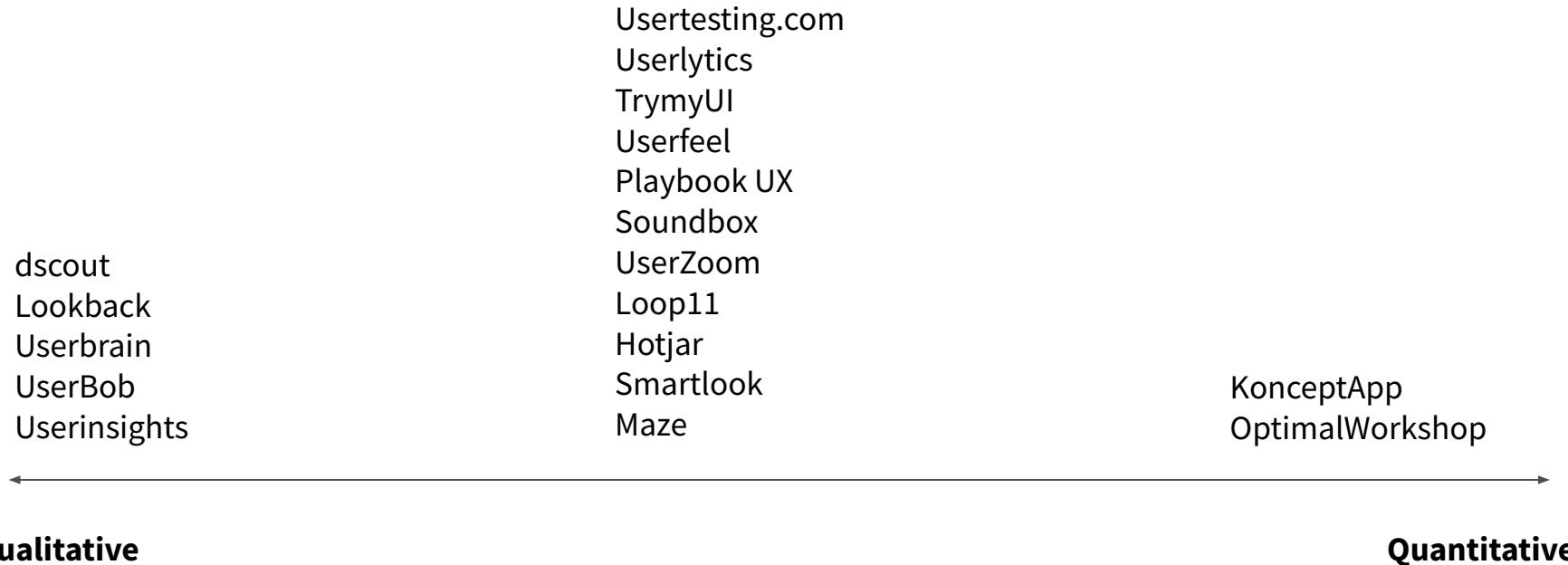
- Record video and audio of user interactions on a native mobile app
- Ask background and follow-up questions
- Tag interactions, create clips and compilations videos

## 2. Select the software

**Based on the goals that you have and your study requirements, select your software.**

- Avoid having too many providers—it can get expensive and confusing
- Consider your goals for potential future projects ahead of time
- It's hard to move from one platform to another

# Lots of unmoderated software out there



# Feature considerations

Recruiting	Study Design & Setup	Qualitative Data	Quantitative Data
<ul style="list-style-type: none"><li>• Participant panel</li><li>• Set quotas for multiple types of users</li><li>• Custom screening question(s)</li><li>• Multiple Languages</li><li>• Bring your own users</li><li>• External panel integration</li></ul>	<ul style="list-style-type: none"><li>• Platforms (desktop, mobile, tablet)</li><li>• Test native mobile apps</li><li>• Test static wireframes or screens</li><li>• Separate instructions for each task</li><li>• Persistent access to task instructions</li><li>• Custom welcome &amp; final screens</li><li>• Copy a past study</li><li>• Branching skip logic</li><li>• Randomize task order</li><li>• Professional research services available</li></ul>	<ul style="list-style-type: none"><li>• Record screen &amp; audio</li><li>• Record face</li><li>• Data tagging tools</li><li>• Timestamped notes</li><li>• Export notes</li><li>• Download recordings</li><li>• Download entire project</li><li>• Share recordings via url</li><li>• Produce video highlights compilation</li><li>• Automatic transcription</li><li>• Browse video thumbnails</li></ul>	<ul style="list-style-type: none"><li>• Simple rating questions</li><li>• Custom ratings and written questions</li><li>• Task time</li><li>• Filter out speeders and cheaters</li><li>• Rate of task abandonment</li><li>• Data export (csv or xls)</li><li>• Data-visualization charts</li><li>• Success rate</li><li>• Click heatmaps and clickpaths</li><li>• Other types of testing</li></ul>

# 3. Create the tasks

## Just like what we learned for moderated testing

- Start broad and then go narrow
  - Task 1: Explore the page and tell me what you see
- Make tasks realistic
  - Structured around your learning goals for the study
  - Specific without being so specific that they overconstrain the situation so that the user goes on auto-pilot
  - Example:
    - Learning goal: Understand if users can successfully organize their photos
    - Poor task: Take every photo of a dog that you took on Thursday and organize it into a folder called dogs.
    - Good task: *Organize the photos that were taken last week in any way that makes sense to you.*

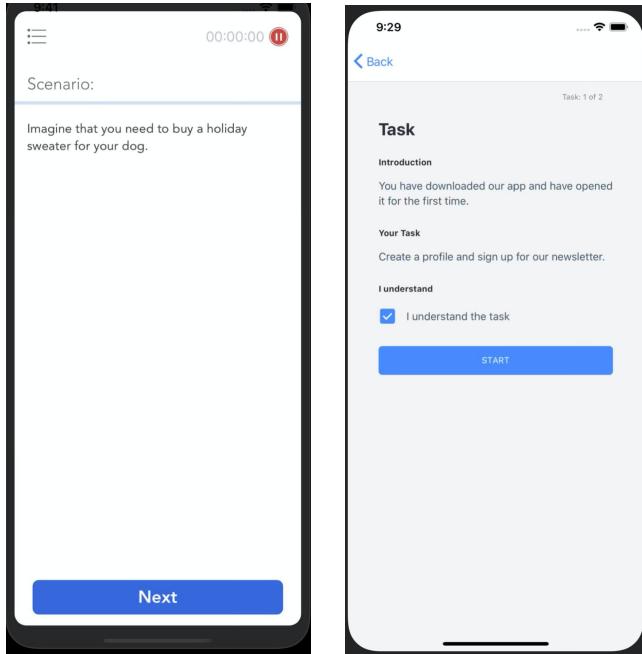
### 3. Create the tasks

- Make tasks actionable, not a set of instructions
  - Ask users to do specific actions without instructing them what to press to do the action
  - Avoid words that are in the interface in your task; instead describe what you want people to do without biasing them towards your language.
  - Example:
    - Learning goal: Understand if users can successfully organize their photos into folders.
    - Poor task: You want to organize your photos into folders. Go to the photos page and show me where you would click New folder.
    - Poor task: Let's organize your photos. Go to the photos page, select some photos, and then click Add to folder.
    - Good task: *Use the Photos page to organize your photos into groups of your choice.*

### 3. Create the tasks

- Be very direct
  - Ask users to do something directly
  - Do not describe situations in general or ask users to “Tell you” something in the task
  - If you need to give specific information to make your prototype work, add it after the instructions
  - Example:
    - Poor task: You own a border collie and a labradoodle. You’d like to have your dog photos organized by breed.
    - Good task: “You’d like to organize your dog photos into groups for each breed of dog. Using the following information, group together your photos for the two kinds of dogs you have. STOP once you’ve created those groups.
      - Dog breed 1: Border collie
      - Dog breed 2: Labradoodle

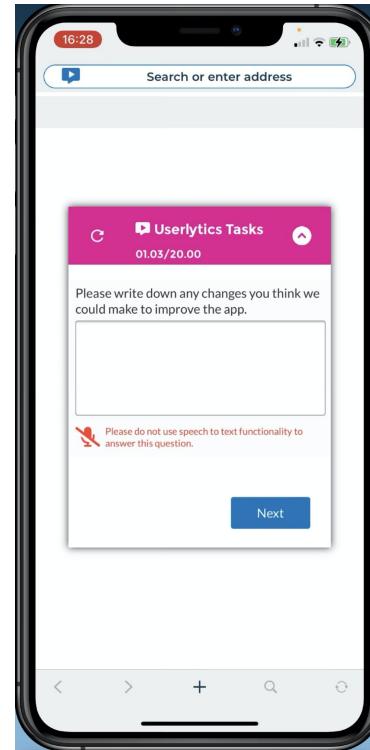
# Keep in mind how users will see the tasks



- Tasks must be short and memorable – do only **one** task at a time
- In a mobile environment, tasks will not be on screen while people are completing them
- Keep instructions to a minimum and use simple, jargon-free language
- If they need to look at a specific area on a screen, describe it with multiple descriptors “Red circle at bottom of page” NOT “Red circle”

# 4. Create strong follow-up questions

- For each task, add **1-3 follow up questions** after the task is completed
- Choose **verbal over written responses** when possible: Verbal recorded responses tend to be higher quality and can be transcribed
- **Avoid leading questions**—especially in these unmoderated studies
  - Leading questions have yes/no answers
- Be simple



# Are multiple choice survey questions OK?

**Survey questions are tempting, but often unreliable**

- **Rating difficulty (1-7):** people often assume they are bad at things and discount their errors
  - After asking for a rating, follow up with the question “What would it take to get it to the next level?” which is where you can actually gather interesting data
- **Best for overall attitudes and perceptions**
  - Example: Rating interest in content can be a good overall indicator
- Can also be used to **see what they remembered**

## 5. Add polish

- Just like in moderated studies, start with a **2-3 background questions**
  - What do you use to take photos?
  - Tell me about the last photo that you shared on social media.
- Create **3-5 wrap-up questions** at the end
  - How did it go?
  - What was confusing about the experience?
- Add an introduction with context about the overall scenario, the fidelity of the prototype, and any other info a participant may need
- Make instructions easy to scan with bullets, bolding, caps and numbering

# Ship it!

- Put it all together into a test plan
- Check if the service you are using offers a free double check of your plan
- Use the service software to input your test plan into their format
- **Run a pilot, run a pilot, run a pilot**



# Synthesizing results: What's different?

- **Start with quantitative data**
  - Understand where the fires are
- **Review the videos** to try and understand WHY the fires are happening
  - Go back to the first video after watching all of them—you'll always spot something new
  - Many services offer tools to help code videos
- Look for not just what people say or do, but **what they don't do**
  - What did they miss? Scroll past? Not react to?
- **Nits & Needs** works well here too!



# Key Takeaways

1. **Unmoderated User Testing** is a method for testing highly functional digital products and services using online software to **understand usability and workflow issues** with a pre-planned script and no moderator.
2. Although they can be run quickly and can even yield quantitative results, they have a **limited depth of insight** compared to moderated testing.
3. To run an unmoderated user test, you need to plan for the following **key elements**:
  - **Prototype / Live site or app:** The artifact being tested must be able to run on its own.
  - **Participants:** Unmoderated service providers can help recruit consumers. For enterprise users, you usually have to provide the participants.
  - **Test plan:** A test plan with scenarios and tasks are key for an unmoderated study.
  - **Software service:** It's important to pick the right testing provider based on your test plan.
4. **Test plans** for unmoderated tests **must have carefully crafted instructions** for users that are actionable, short and memorable. Users often can't see instructions while completing a task on mobile.
5. Understanding unmoderated test results involves **watching and coding videos** in addition to any quantitative data evaluation.

# Summary

# Course Recap

## 1: User Research Methods

- Overview
- Introduction to User Research
- Types of Methods and their Attributes
- User Research within the Product Development Process

## 2: Deep Dive: Rapid Experimentation

- What is Rapid Experimentation?
- How to Run a Rapid Experiment
- Subject Recruiting
- Synthesizing and Sharing Results

## 3: Deep Dive: Moderated User Testing

- What is Moderated User Testing?
- Test Moderation
- Creating a Moderated Test Plan
- Synthesizing and Sharing Results

## 4: Deep Dive: Unmoderated User Testing

- What is Unmoderated User Testing?
- Creating an Unmoderated Test Plan
- Synthesizing and Sharing results

# Key Takeaways

1. User research is a great tool for **understanding your customers** and **gathering their feedback on your ideas** throughout the product development process.
2. Three key, effective methods that are widely used are: **rapid experimentation**, **moderated user testing** and **unmoderated user testing**
  - **Rapid experimentation** is a method that helps product managers **understand how people might actually engage with a new idea** in the real world before it actually exists. Planning is a five step process that involves carefully designing an experiential prototype to answer a specific question about your idea.
  - **Moderated User Testing** is a method for gathering feedback on design ideas and prototypes where moderators guide participants to try a prototype and answer follow-up questions about it. Planning involves crafting a test plan of tasks and questions for users and practicing effective moderation techniques.
  - **Unmoderated User Testing** is a method for testing highly functional digital products and services using online software to **understand usability and workflow issues** with a pre-planned script and no moderator. Planning involves writing very clear instructions for users and selecting a service with features that support your test plan.
3. Choosing which method to use depends on your **stage in the product development** process, the **questions** you are answering, the **state of your prototype** and how much **time** and **resources** you have available
4. Any user research is better than no research so of the the best type of testing is **the kind you actually do!**