Evaluation: Usability testing

CPSC 544: Fundamentals in Designing Interactive Computational Technology for People

Material created by Joanna McGrenere, Karon MacLean, Leila Aflatoony, Jessica Dawson & Heather O'Brien

Today

- Announcements
- Usability testing lecture
- In class activity [35 min]: Plan a usability study

Coming up

This week (Oct 30)

- Sun 10/29:
  - Researcher Journal #11
- Tues 10/31
  - Researcher Journal #12

- DO: if you haven't already, meet with your staff mentor early this week to discuss low-fi prototyping strategy (Thu lab is a bit late)

Coming Up

- Mon 11/06: Low-Fi Prototype Presentations
- Wed 11/08: In-class: Cog Walkthrough of lo-fi prototypes

- Reading Week: TA mentors can meet 1x with each team, schedule in advance

- Tues 11/21: User Test Report Due (soon after break)

Heads-up:
Learning Self-Reflection

- Reflect on what you learned over the term
- Due Dec 10 (same day as final report; full instructions on ISched, Dec 10)

End-Term Learning Self-Reflection Assignment

Total 2 page max

We invite you to reflect on what you've learned, by identifying **three points where something notable happened for you**, in the context of **any aspect of the course**, For example, class or readings, project content or teamwork itself are all reasonable things to consider, as well as others.

We also encourage you to (privately, not here unless it is related) reflect on how you participated and contributed to your team and the class.

Please include in your reflection how the point was notable: e.g., was it a **surprise**, a **challenge** (where something got tough, and did or did not ultimately work out); something you were **especially proud of**, or something you'd **do differently next time**? Or something else.

For **each** of these three points on next page (please don't alter the template formatting):

1. Provide sufficient detail for the teaching team to understand the **context of the point/event/situation you are describing**. E.g., Which project milestone / paper reading, and within that, additional detail to set up your account. *[1-2 sentences usually enough]*
2. What made the point notable in the way you have identified (e.g., surprise, challenge etc). *[1-3 words]*
3. Explain **what you learned as a result of this point / event**: e.g., about specific topics covered in the course or HCI design methods and techniques, working on a team, etc. *[Short paragraph]*
4. Note how this was **significant for you personally.** E.g., did you take risks or stretch yourself? Did you have to change the way you typically engage in courses due to the collaborative nature of the course? Did you make a difference for others on your team? How do you think that the thing you learned will matter for you going forward? *[Short paragraph]*

DFP Seminar – Nov 8.
Don't forget to RSVP

Non-Contact Monitoring of Respiratory Rate for Infants in the NICU using Depth and RGB Cameras

Paul Addison and Soodeh Ahani

Paul Addison, PhD, Chief Scientist Medtronic Patient Monitoring, Technical Lead Algorithm and AI Group, Professor of Fluids Engineering | Soodeh Ahani, PhD, Research Associate, UBC

November 8th, 2023, 12:00pm

2023 Series

DFP Seminar

Join us on November 8th as Dr. Paul Addison and Dr. Soodeh Ahani explore "Non-Contact Monitoring of Respiratory Rate for Infants in the NICU using Depth and RGB Cameras".

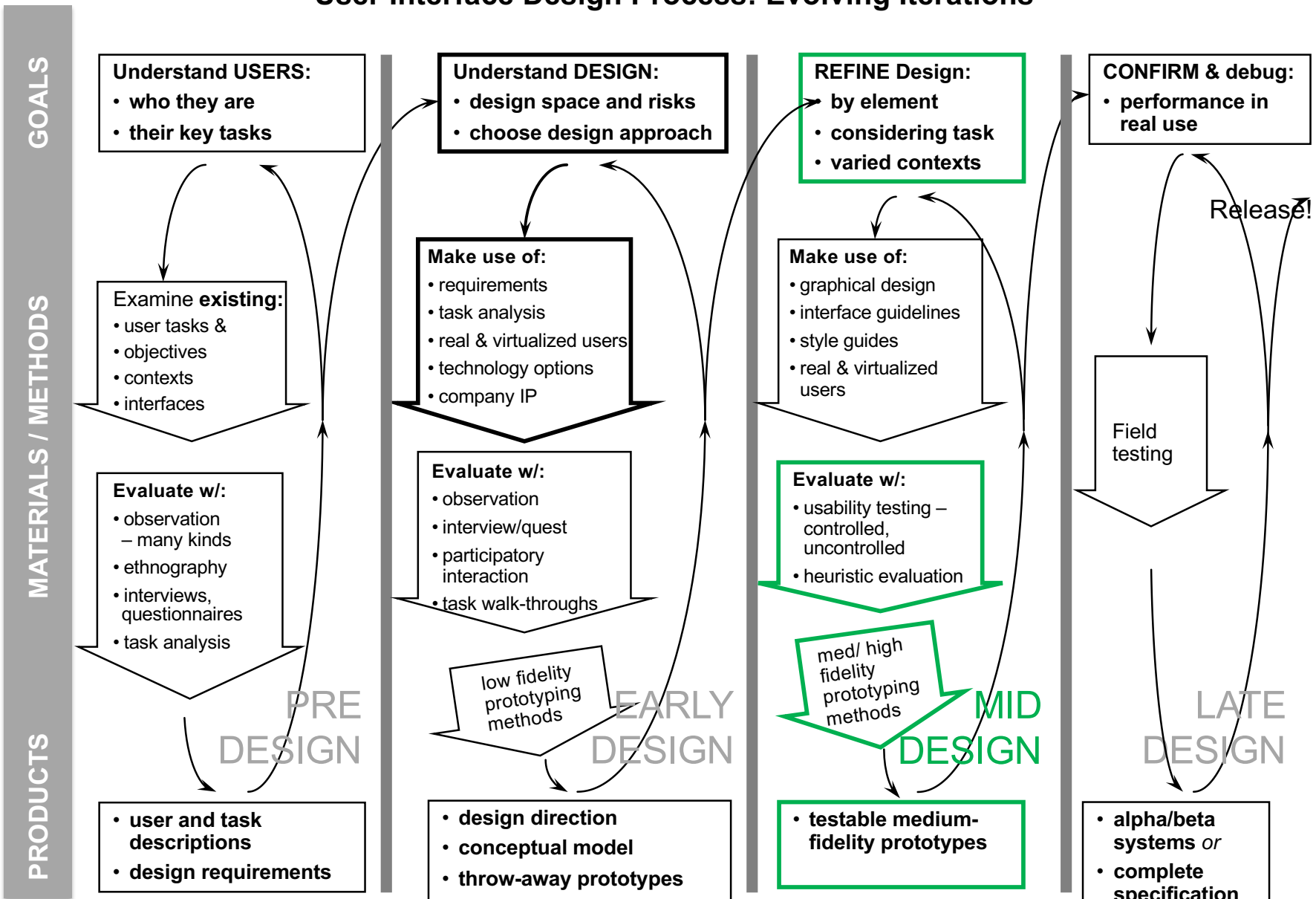
RSVP here!

https://dfp.ubc.ca/news-and-events/events/dfpubc-seminar

Learning Goals

- Understand the role of usability testing in HCI
- Distinguish usability testing from other evaluation methods
- Define usability and usability testing
- Explain when usability studies are typically conducted and why
  - Give examples of locations, tasks, metrics, evaluation methods that might be involved
- Describe how to plan and conduct a usability study

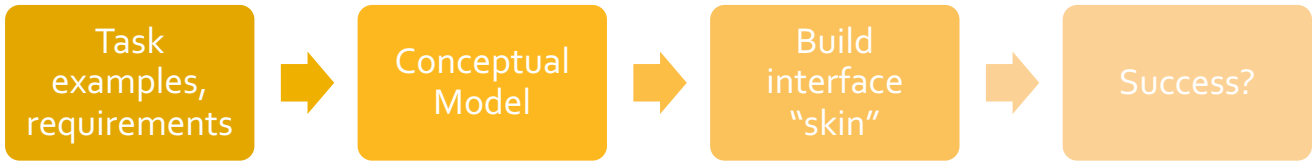
User Interface Design Process: Evolving Iterations



K MacLean - derived from version by Saul Greenberg (U Calgary)

What is the role of usability in HCI?

- Usability: a primary focus of HCI
  - **Evaluate** system usability
    - how easy it is for the user to get the system to do what they need it to do?
  - **Design** for usability
  - **Establish/apply** metrics and standards for usability



Note on terminology

- Not entirely standardized…
  - **User Study** – very general term. Any study that involves actual or prospective users.
    ***Timing**: In different forms, **anytime** -- from before a system is built (Empathize / Pre-Design) up to a controlled experiment with a field-deployed system.*
  - **Usability Study** – more specific. Requires a system for which task performance can be measured
    ***Timing**: usually Mid / Late Design, but can be Pre-Design for a system being re-designed. Today's focus*
  - **Controlled Experiment** – a specifc type of usabilty study with hypotheses and statistical testing, often comparing alternate designs (more on this later).
    Timing: Test / Late Design. Upcoming lectures
  - **Informal / Small User Study** – often used before a usability study, not ready to measure things yet, interested in higher-level feedback.
    ***Timing**: Early design. You've already done one of these*

How is usability testing different from other evaluations?

Heuristic evaluation

Design team

Find major issues using set of heuristics

Cognitive walkthrough

Colleagues

Find issues re: walkthrough of tasks

Usability testing

"Real" users

Utilize tasks BUT include performance and perception metrics

May draw on multiple methods

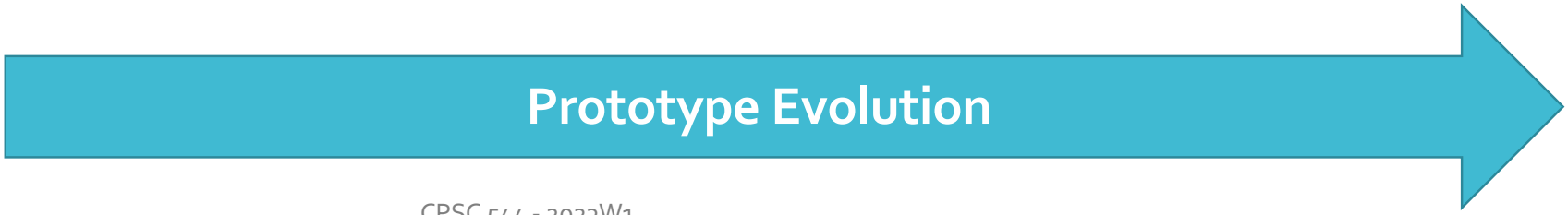
One system

Experiment

"Real" users

May focus on multiple systems (A/B testing)

May focus on different populations or conditions



Alternatives/ complements to usability testing

Earlier this term

Last class

- Observations, surveys, interviews also great in field studies
  - Evaluating to understand: often looking for qualitative insights
    - stories, workflows, obstacles, dependencies, missing links.
- "Discount" methods can also target prototypes at various stages and be done without users
    - heuristic evaluation
    - cognitive walkthrough
  - Because you don't need users . . .
    - can do it first (before a usability study)
    - possible to apply these methods yourself while iterating on a design (before it's totally finished)
- Usability testing **requires users, relatively refined prototypes**, and usually focusses on **measuring something**
  - Testing for performance, learning challenges, acceptance …

Biggest differences with alternatives:

- Usability testing requires:
  - A refined interface.
  - This could be… your new medium fidelity prototype.
  - Or it could be the bad old interface, which you plan to revise or replace
    - i.e., might be "evaluate for understanding the problem"
- Measured outcomes
- Users (participants)

How do Greenberg & Buxton position **usefulness** and **usability**?

What is usability?

Utility = whether it provides the **features you need**
(and the features work, in a technical sense)

Usability = how **easy** + **pleasant** the features are to use.

Useful = **usability + utility**
(does the job, in an easy, pleasant way)

Nielsen, 2012

Aspects of Usability (Nielsen, 2012)

- **Learnability**: easy to learn so a user can rapidly start to use it
- **Efficiency**: once the user has learned the system, a high degree of productivity is possible *(aka performance)*
- **Memorability**: the user should be able to return to the system and not have to learn again
- **Errors**: users should make few errors and recover easily *(aka, effectiveness)*
- **Satisfaction**: the system should be pleasant to use

Usability study/test:
→ Evaluates an interactive system or prototype with respect to all/some of these elements,
always involving real users

Elements of a usability test

- Interactive system / prototype
- Evaluation goals
- Tasks
- Measures/metrics
- Data collection/recording methods
- Participants

When designing a usability test:

- **Choice of methods:** triangulate
  - Typically: one instrument counts something, while another interprets what was counted
- **Choice of metrics:**
  - Driven by your requirements, evaluation goals, and basic usability principles
- **Which/how many users?**
  - Should be representative of your user groups *(may need to approximate "ideal" user)*
    - e.g.: if you want to support both expert and novice users, should have good numbers of both!
  - Much debate in the literature on the "right" number (Lazar et al.)
  - Sometimes constraints dictate low numbers
    - Examples?
    - If you have to generalize, consider who your test users are, and how representative they are?

Tasks

- Generally: Task specification arises from user research
- Can be:
  - Low level: e.g. the subtask that will take you from one screen to the next.
  - Entire task level: see if someone can figure it out, start to finish, and watch /count / measure the challenges s/he has
- Not done with those task examples and task descriptions yet!

Methods

- **Observational techniques:**
  - Silent
  - Think aloud (participant talks as they do the task)
    - Constructive interaction (participant is probed periodically)
  - Think after (Interpretation session)
- **Query techniques:**
  - Interview
  - Survey
  - Questionnaire

"…how we can precisely distinguish **design flaws** caused by usability issues or **utility problems** with the data from usability testing"

Metrics

- **Time: to**
  - Complete a task (entire, or a portion)
  - Learn a task
  - Resume a task after interruption
  - Find something on a screen
  - Attain specified degree of proficiency
- **Errors:**
  - Number per task or unit of time
    - different types of "number": e.g., what do you count for errors in navigation, selection, interpretation?
  - Number of users making the error
  - Locations of errors (as well as numbers)
  - Alternately, number of successes

Metrics (continued)

- **Events of interest:**
  - page views or clicks
  - access of particular tools
  - timeouts
  - questions asked or help tools consulted
  - # users willing to recommend
- **Subjective factors:**
  - task level satisfaction
  - perception of aesthetics
  - perceived ease of use
  - perceived preference
    - Can be measured on a Likert or semantic rating scale

+ standard questionnaires for Usability and User experience

Example of a usability test: Help Kiosk (HK)

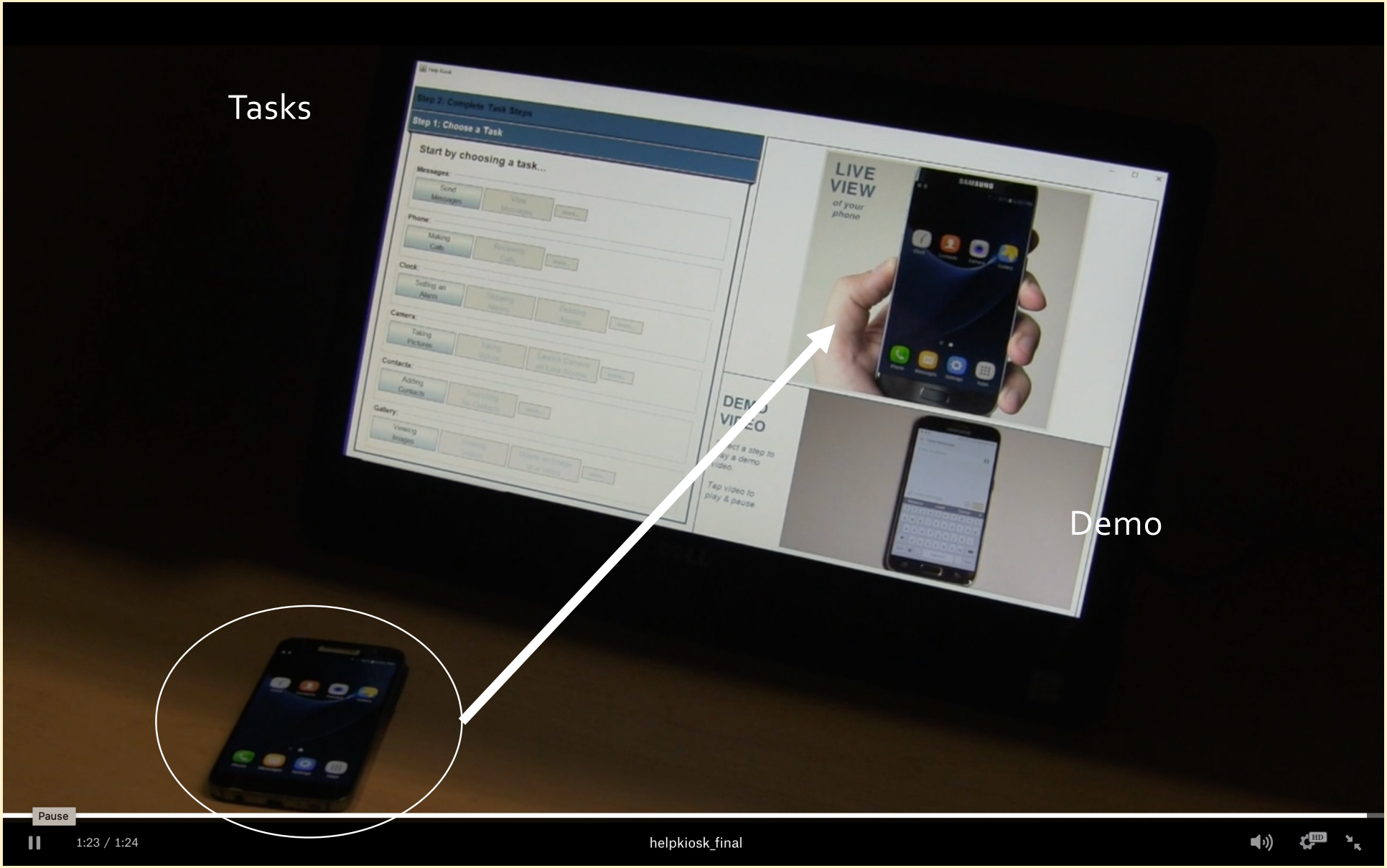
An Augmented Display System to Assist Older Adults to Learn How to Use Smart Phones *(2018 DFP Project)*

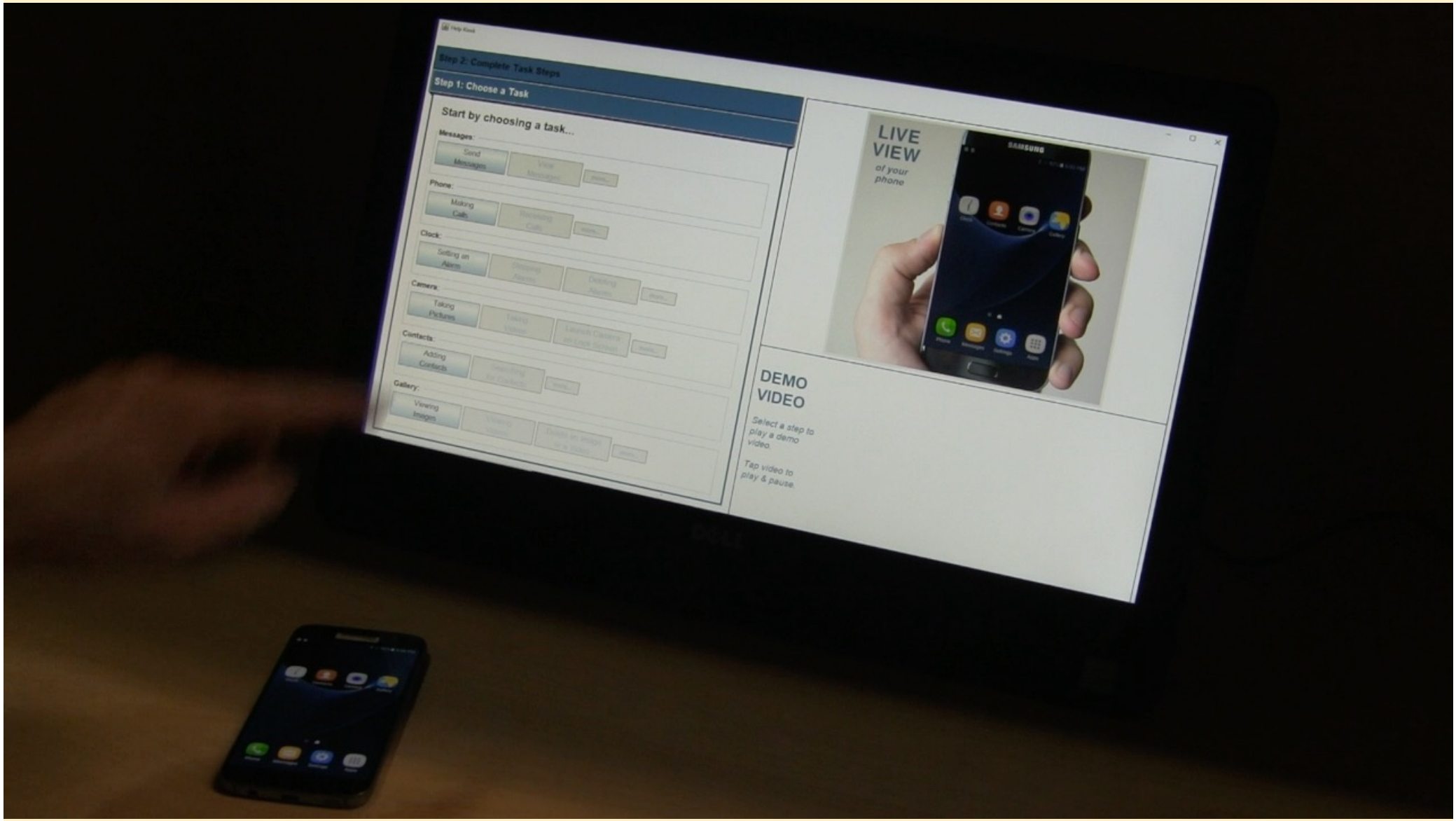
Lead: *Joanna McGrenere*

Collaborators: *Sayan Sacar (University of Tsukuba), Rock Leung (Samsung), Karyn Moffatt (McGill), Peter Graf (UBC), Helen Yin, Zachary Wilson*

Source: https://dfp.ubc.ca/project/help-kiosk

- **Evaluation goals**
  - To see if older adults can learn to do basic tasks independently on their own smart phone after using HK for a short period of time (<30 min)
  - To see if the concept of Help Kiosk is appealing and would be preferred to using a manual or getting help from a friend
- **Tasks**
  - Use HK to learn how to
    (1) a contact to your smartphone,
    (2) send a text message, and
    (3) call a contact
  - Repeat the same tasks without the support of HK, using printed manual


Tasks
Step 2: Complete Task Steps
Step 1: Choose a Task
Start by choosing a task...
Messages:
Send Messages
Phone:
Making Calls
Clock:
Setting an Alarm
Camera:
Taking Pictures
Contacts:
Adding Contacts
Gallery:
Viewing Images
LIVE VIEW
of your phone
DEMO VIDEO
Tap video to play & pause
Demo
Pause
1:23 / 1:24
helpkiosk_final


Step 2: Complete Task Steps
Step 1: Choose a Task
Start by choosing a task...
Messages:
Send Messages
Phone:
Making Calls
Clock:
Setting an Alarm
Camera:
Taking Pictures
Contacts:
Adding Contacts
Gallery:
Viewing Images
LIVE VIEW
of your phone
DEMO VIDEO
Select a step to play a demo video.
Tap video to play & pause.

DEMO (90 sec)

Example (cont'd)

- **Measures/metrics**
  - Time: from the moment user starts until they complete each task successfully
  - Errors: number of taps that deviate from correct task path
  - Satisfaction: Likert scale
- **Data collection methods**
  - Observation - video recording
  - Observation – software logging of all interactions
  - Survey – to capture satisfaction
  - Interview – probe on user experience and comparison to other learning methods
- **Participants**
  - 16 older adults (60+), mix of genders, range of computer expertise, all new to smartphones

Usability testing in your **project**

- **What design stage does a usability study fit into?**
- **1. What are your evaluation goals? First step!!**
  - Draw from your requirements and task examples; may need to prioritize.
  - Goal: test how well your system supports what you intended it to (that is your "research question")

  → 2. What **Metrics** will **address your eval goals?**

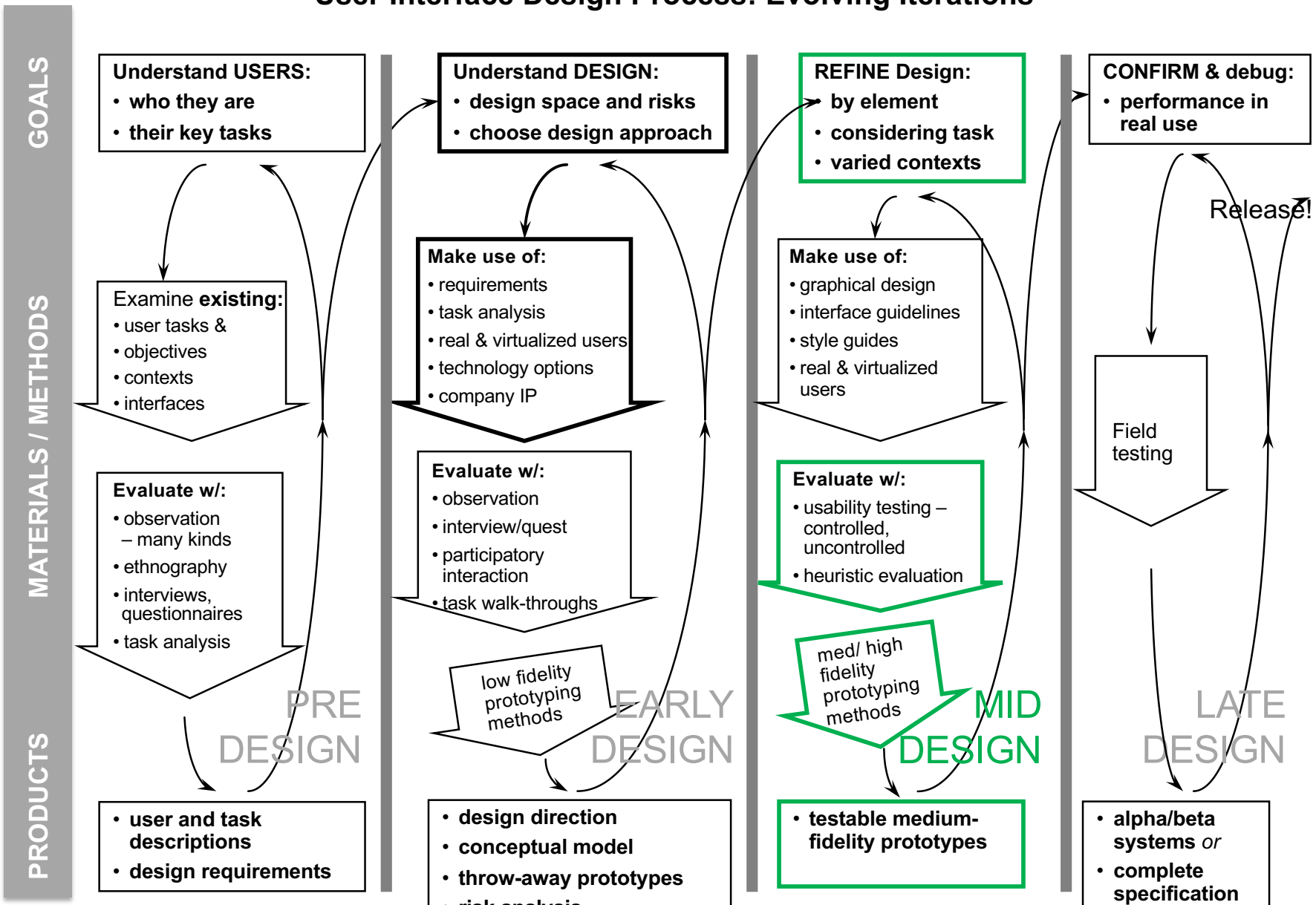
  → 3. What **Evaluation methods** will deliver those metrics?

  → 4. What **Prototype** will support that evaluation?

- **Medium fidelity prototype scope?**
  - Prototype won't be a complete working system
  - It should do just enough to test if your design will meet your goals (and be achievable in the time available)

You'll need those informed consent templates again!

User Interface Design Process: Evolving Iterations



K MacLean - derived from version by Saul Greenberg (U Calgary)

In-Class Activity: **Plan a Usability Study**

up to 35m

INSTRUCTIONS [5min]

- Each team split up:
  - <= half: "**Sharers**" (1/3 or 2/4 members)
  - >= half: "**Consultants**" (2/3 or 2/4 members)
- **Sharers** join another team's **Consultants** (see swap plan on next slide)
- **[5min] Sharers:** describe **your project** to the Consultants you're with.
  - Focus on the more unusual, high-risk aspects – the parts that will need testing.
  - Use visuals from your laptop as needed (e.g., conceptual model, task descriptions).
- **[up to 20m] Sharers + Consultants:** discuss the usablity planning questions on upcoming slide
  - Recommend **goals, tasks metrics, methods**.
- **Sharers, take notes!**
- **[5min] Original teams reconvene**.
  - Sharers share feedback with your team members who have been consulting for another team.

In-Class Activity: **Swap Plan**

Sharers from …	work with **Consultants** from:	At spot
Dating for Wellness	Smart Insole	1
Smart Insole	Peer Learning	2
Peer Learning	Campus Security	3
Campus Security	NLP for COPD	4
NLP for COPD	Dating for Wellness	5

In-Class Activity: **Plan a Usability Study**

Reminder: what design stage will this study be in?

Get back into Project teams, and DISCUSS:

- Sharers, first give your teammates a quick synopsis of the feedback you received.
- What would be good evaluation **goals**?
  - *Think about: why are you doing this study, given design stage?*
- What would be a good **study task**? (for this evaluation)
- What are good **metrics** for that task?
- What **evaluation type** and **data recording methods** would be best for creating and collecting those metrics?

References

- Greenberg, S. & Buxton, B. (2008). Usability Evaluation Considered Harmful (Some of the Time). In Proceedings of the 2008 CHI Conference on Human Factors in Computing Systems (CHI '08). ACM, New York, NY, USA, 111-120.
- Lazar, J., Feng, J. H., & Hochheiser, H. (2017). Usability testing. Analyzing qualitative data. Research methods in human-computer interaction. (263-298). Morgan Kaufmann.
- Lewis, C., & Rieman, J. (1993). Testing the design with users. Task-centered user interface design. A practical introduction (pp. 77-96).
- Nielsen, J. (2012, January 3). Usability 101: Introduction to usability. NNg, Retrieved, http://www.nngroup.com/articles/usability-101-introduction-to-usability/