Think-Aloud Usability Evaluations

SWE 632
Spring 2018

Administrivia
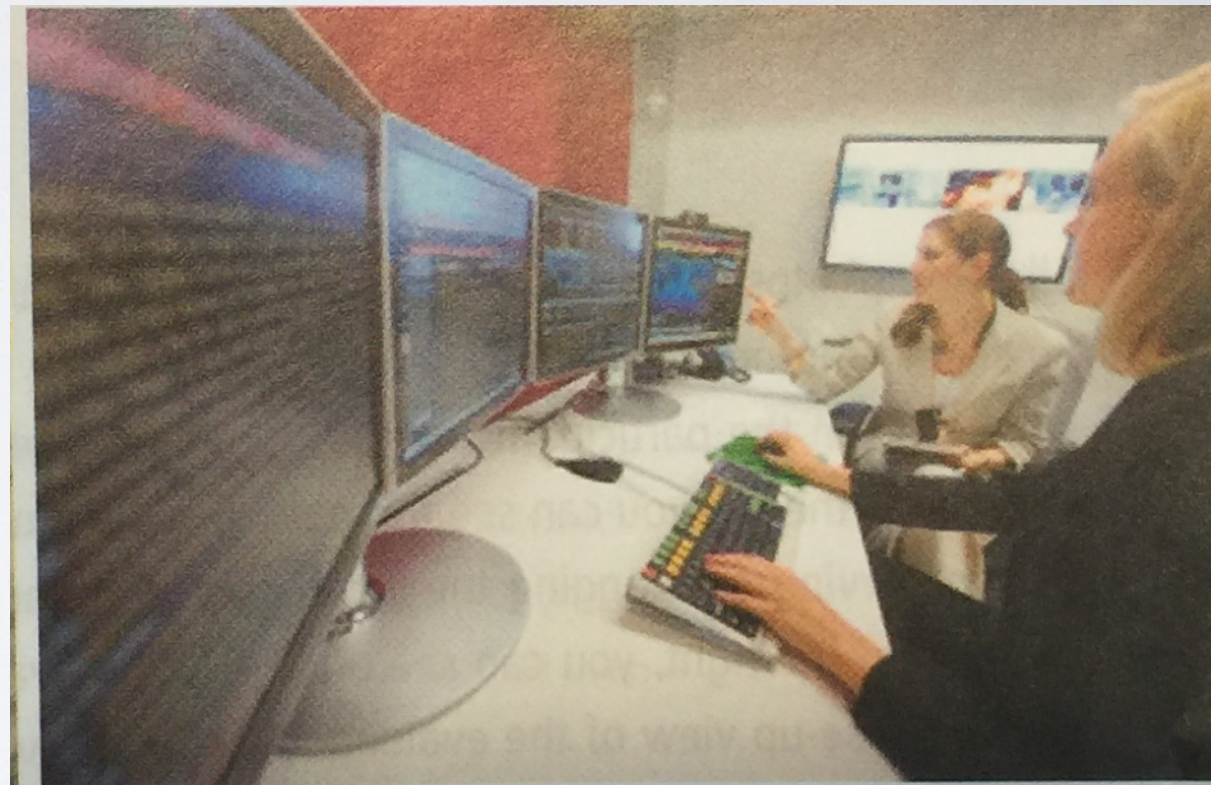
- HW 3 due 3/1 (Thursday)

- In-class midterm next week

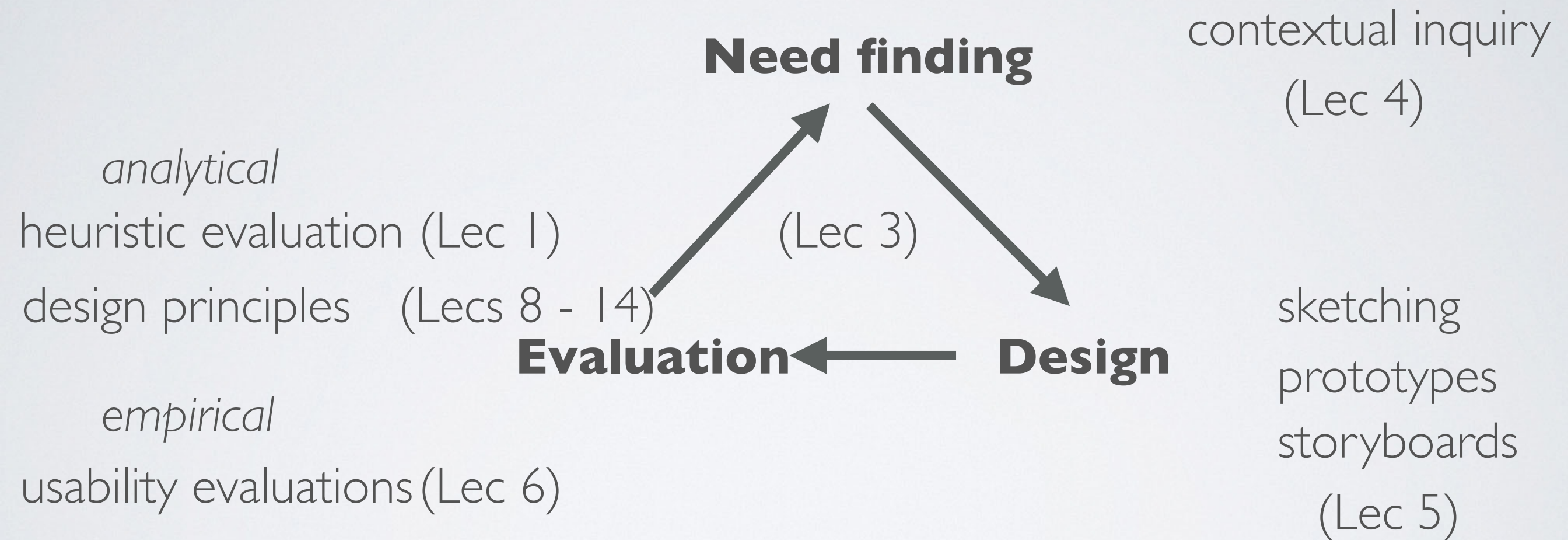- HW 4 out next week, due 3/27 (4 weeks from today)

Midterm review

Usability Studies

Why conduct usability studies

- Evaluate interaction design with **real** empirical data, gathering ground truth of user performance

- Identify **usability issues**

Design Process Big Picture

Need finding

contextual inquiry
(Lec 4)

analytical

heuristic evaluation (Lec 1)

(Lec 3)

design principles   (Lecs 8 - 14)

Evaluation ← **Design**

sketching

empirical

prototypes

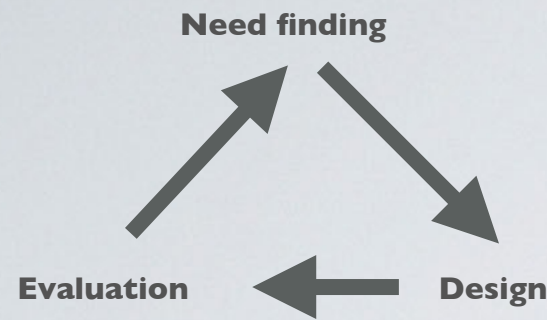usability evaluations (Lec 6)

storyboards
(Lec 5)

Think-aloud usability study

- Goal: observe users using app, identify usability issues

- Can use with

  - paper prototype

  - HTML prototype

  - Wizard of Oz study

  - actual app

Steps in a usability evaluation study

- Formulate **goals** of study

- Design study protocol, tasks, materials, data collection, …

  - Pilot study design

- **Conduct** study

- **Analyze** data to assess task performance and identify usability issues

Formulate study goals

Need finding

Evaluation　　**Design**

Study goals

- Where are you in the design process? What feedback do you seek?

  - Exploring new design idea

  - Validating high-level approach

  - Identifying important usability issues

  - Evaluating a new feature just added or a particular corner case

  - Studying performance by specific users (e.g., expert users familiar with old version)

  - Comparing performance against competitors

Study design

Selecting participant population

- Who will be the users?

- Goal: users representative of system's **target users**

- Are there multiple **classes** of users (e.g., data analysts, site administrators)?

  - If so, which are appropriate given goals?

  - May choose several classes

- System **novices** or **experts**?

- Might choose to include **UX experts** to help flag potential issues

Number of participants

- More participants —> different participant interactions, more data

- Fewer participants —> faster, cheaper

- No right answer, as depends on potential diversity of interactions and users

- Nielsen & Morlich (1990) found that 80% of problems could be detected w/ **4-5** participants

  - Most serious usually detected with first few

  - Krug suggests 3

Informed Consent

- Important for participants to be told up front what they will do and provide affirmative consent

- Helps allay potential participant fears

- Make clear purpose of study

- Make clear that you are evaluating your design, **not** the user

Tasks

- What will users do?

- Goals for task design:

  - Provide specific goal: something that the user should accomplish

  - Comprehensive enough to exercise key features of your app

  - Short enough to minimize participant time commitments

Communicating tasks

- Provide a scenario explaining the background of what users will be doing

- Provide a specific goal that the user should accomplish

  - But **not** how they should accomplish it

  - Don't give away how you hope users will accomplish goal

- Communicate **end criterion** for task - how do they know they're done?

- Provide maximum time limit after which they will be stopped

Recruiting participants

- Many potential sources

  - Co-workers, colleagues, friends, family

  - Email, mailing lists, online forums

  - Announcement at related user groups

- Important to select sources that best match the background & knowledge of target users

Incentives for participants

- Often (but not always) helpful to pay participants

- Most applicable when seeking participants with specialized expertise with whom you do not already have a personal or professional relationship

- Can also offer other incentives, such as gifts, coffee mugs, gift certificate; or free consulting, training, or software

- In some cases, just learning about future product can be incentive

Managing participants

- Participants are valuable resource

  - Often finite resource

- Think carefully about how participants will be used

- Devise mechanisms for scheduling participants & reminders

Training

- Goal: **avoid** unless really necessary

- Training necessary when

  - Participants require specialized knowledge to act as target users

  - Target users will have access to specialized training materials before they begin study

Data collection

- Think aloud

- Screencast

- Questionnaires interview questions to gather participant feedback

Questionnaires and interviews

- Gather background or demographics about participants (if important)

- Supplement task performance data with subjective reactions

  - Perceptions of design, comments on potential issues, ideas for features

- Questionnaire - pre-defined questions, focused, less bias

- Interviews - more open ended, longer responses

Example open-ended questions

- What did you like best about the UI?

- What did you find most difficult or challenging?

- How might the UI better support what you're trying to do?

Piloting study design

- Dress rehearsal for conducting actual study

- Goals

  - Ensure software / prototype won't "blow up"

  - Test tasks - ensure right length & difficulty

  - Test that materials are comprehensive and comprehensible

- As-needed piloting

  - Use first study session as pilot only if issues arise and must be addressed

Conducting the study

Introduction (1)

- Greet participants, introduce yourself, thank them

- Build rapport, socialize

- Introduce them to the setup

Introduction (2)

- Give participant Informed Consent

- Answer any questions about study design

- Relieve anxiety and curiosity as much as possible

- Make clear evaluating design, not participant

- Let participants know you can't answer questions about how to do task

Starting session

- Give participants description of task

- Start any video recording

- Start encouraging participant to think aloud

- Begin observing participants work on task

Interactions during the task

- Goal: listen, not talk

- Prompt participants to think aloud when necessary

  - e.g., What are you trying to do? What did you expect to happen?

- If show signs of stress / fatigue, let them take a break

- Keep participants at ease

  - If participants frustrated, reassure & calm participants

  - If so frustrated they want to quit, let them

Giving help

- If participants totally off track, small reminder of goal might help

- Should **not** give participants information about how to complete the task

- What if user asks for help?

  - Direct them to think through it or work it out for themselves

Collecting critical incidents

- *Any action that does not lead to progress in performing the desired task*

- Often related to a gulf of execution or gulf of evaluation

- Generally does not include

    - accessing help

    - random acts of curiosity or exploration

Understanding a critical incident

- Important to understand in the moment what users goal is and what actions they are taking

- When a critical incident occurs, jot down

  - The time

  - What user was trying to do

  - What user did

Wrapping up the study session

- Provide questionnaire (if applicable) / conduct interview (if applicable)

  - Probing into causes of behavior

- Answer any lingering questions the participant may have

- Thank the participant!!

- Provide any incentives (if applicable)

Reset study environment

- Make sure study environment is in the same state for all participants

  - Reset browser history / cache (if applicable)

  - Delete any user created content or materials

Analyzing data

Critical incident analysis

- Identify critical incidents where something went wrong

- Easiest to catch in the moment - important to take good notes

- Going back and looking at screencast can help you study context of issue in more detail

Reporting a critical incident

- Problem statement: summary of problem and effect on user (but not a solution!)

- User goals: what was user trying to do?

- Immediate intention: at the moment in time when problem occurred, what was the user trying to do

- Possible causes: speculate on what might have led user to take action they did

Critical incidents --> usability issues

- Group together similar incidents to form **usability issue**

  - Match similar critical incidents within and across study sessions

  - Identify underlying cause

- Brainstorm potential fixes

In class activity

Group activity

- Form groups of two (not anyone you've worked with before)

- Take turns conducting a usability study of your project app

  - 5 mins to brainstorm 5-10 min task

  - 15 mins to conduct study

  - Identify critical incidents (if any)