Think-Aloud Usability Evaluations

SWE 632
Spring 2022
Administrivia

• In-class activity policies
• Midterm exam next week
  • Covers all lectures & readings before exam
• No class in 2 weeks (3/16) - Spring Recess
• HW4 will be released next Wed (3/9)
  • Think-aloud usability eval of project app
  • Due in 4 weeks (3/30)
Think-Aloud Usability Studies
Iterative Model of User-Centered Design

**Observation**
(Re)Define the Problem
Understand User Needs

**Test**
Evaluate what you have built

**Idea Generation**
Brainstorm what to build

**Prototype**
Build
Iterative Model of User-Centered Design

Test

Evaluate what you have built
Why Conduct Usability Studies?

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

• Identify *usability issues*
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
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 and 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
Usability Study vs. Contextual Inquiry

**Usability Study**
- Used for evaluation
- Generally conducted via observation
- Identification and analysis of “critical incidents”
- Intended to identify usability issues

**Contextual Inquiry**
- Used primarily for “needfinding”
- Conducted more like a conversation
- Obtain data about users in their context
- Intended to help in the design phase of a project
10 Minute Break
In-Class Activity
Group Activity

• In groups of two or three
• Take turns conducting a usability study of a different web app of your choice
  • 5 mins to brainstorm 5-10 min task for each app
  • 10-15 mins to conduct each study
  • Identify critical incidents (if any)
• Deliverables (due 11pm today)
  • Your name
  • Name of app your evaluating / short description
  • Description of critical incidents