Evaluation with users

Readings:
Dix 9.4
“Ethics in HCI”, CHI2001 Panel Documentation
Announcements

- Last midterm: July 17 instead of July 20
- July 20: No lecture, meet as needed to discuss projects in CLEA 127.
- 2 coop jobs for Fall 2006:
  - Research-oriented
  - UI design
  - Computer vision, Computer graphics, medical image processing
Today’s topics

- User testing
- Ethics
- Formative evaluation
Where we are now

Design
- Task analysis
- Design heuristics

Evaluate
- Heuristic evaluation
- Today: User testing

Implement
- Prototyping
MS Office 12: redesign issues
Kinds of user tests

- Formative evaluation
  - Laboratory study
  - Finds problems for the next iteration in design
  - Evaluates prototype in lab on chosen tasks
  - Qualitative observations (usability problems)
Kinds of user tests (2)

- **Field study**
  - Finds problems in context
  - Evaluates working implementation, in real context, on real tasks
  - Mostly qualitative observations

- **Very expensive**
  - Norman: “Field studies, user observations, contextual analyses, and all procedures which aim at determining true human needs are still just as important as ever – but they should all be done outside of the product process. This is the information needed to determine what product to build, which projects to fund.”
Kinds of user tests

○ Controlled experiments
  ● Test a hypothesis (e.g. interface X is faster than interface Y)
  ● Evaluates working implementation, in controlled lab environment, on chosen tasks
  ● Mostly quantitative observations (time, error rate, satisfaction)
Ethics of user testing

“Users are human. As HCI professionals we must be sure that our fellow humans perceive their encounter with usability and design professionals as pleasant without sacrificing the accuracy of our results”

CHI2001 Panel Documentation, Ethics in HCI
Ethical issues: Case studies

“You go into the field. You have told your interviewees that their comments to you are private. A senior VP asks you “who did well in the study?” What do you say?”

CHI2001 Panel Documentation, Ethics in HCI
“Participants in a usability test have given written consent to have the session videotaped for internal company purposes. They are not explicitly told that some unknown number of people will be watching the live video from the observation and control rooms. Are the live observers covered by the consent form?”

Carolyn Snyder, CHI2001 Panel Documentation, Ethics in HCI
Dilemmas (cont’d)

- Yes: Consenting to being videotaped includes implicit permission to be observed live.

- No: The participant may trust that any uncomfortable moments captured on the tape will not be shown to others, but with live action there’s no way to prevent this. Also, people in a non-public setting have a right to know exactly who is watching them.
Pressures on a user

- Performance anxiety: “Am I doing it right?”
- Feels like an intelligence test
  “Do these people think I’m dumb for not getting it?”
- Comparing self with other subjects
- Competing with other subjects
10 steps for conducting a user observation

- 1. Introduce yourself.
- 2. Describe the purpose of the observation (in general terms).
  - Stress that you want to find problems in the product
    - You're helping us by trying out this product in its early stages.
    - We're looking for places where the product may be difficult to use.
    - If you have trouble with some of the tasks, it's the product's fault, not yours. Don't feel bad; that's exactly what we're looking for.
    - If we can locate the trouble spots, then we can go back and improve the product.
    - We're testing the product, not you.
More steps…

- 3. Tell the participant that it's okay to quit at any time
  “Although I don't know of any reason for this to happen, if you should become uncomfortable or find this test objectionable in any way, you are free to quit at any time.”

- 4. Talk about the equipment in the room.
  - purpose of each piece of equipment (hardware, software, video camera, microphones, etc.) and how it is used in the test.
More steps…

5. Explain how to “think aloud”
   - All you do is speak your thoughts as you work.
   - If you forget to think aloud, I'll remind you to keep talking.
   - Would you like me to demonstrate?

6. Explain that you cannot provide help
   - As you're working through the tasks, I won't be able to provide help or answer questions. This is because we want to create the most realistic situation possible.
More steps…

- 7. Describe the tasks and introduce the product.
  - Explain what the participant should do and in what order.
  - Give the participant written instructions for the tasks.
  - Don’t demonstrate something you are trying to test.

- 8. Ask if there are any questions before you start. Then begin observation.
Last steps

9. Conclude the observation
   - explain what you were trying to find out during the test.
   - answer any remaining questions the participant may have.
   - discuss any interesting behaviors you would like the participant to explain.

10. Use the results
Roles in formative evaluation

- User
- Facilitator
- Observers
User’s role (1)

- **Simple observation**: just watching the user
  - In lab: user is asked to complete a set of predefined tasks
  - In field study: user goes through normal duties
- **Problem**: does not give insight to the user’s decision process and attitude
User’s role (2)

- Users should **think aloud**:  
  - What they think is happening  
  - What they’re trying to do  
  - Why they took an action

- **Problems:**  
  - Feels weird  
  - Thinking aloud may alter behaviour  
  - Disrupts concentration
User’s role (3)

○ Constructive interaction:
○ Normal conversation between two users is monitored
Facilitator’s role

- Does the briefing
- Provides the tasks
- Coaches the user to think aloud by asking questions
  - General:
    - “What are you thinking?”
    - “Why did you try that?”
  - More specific if user stuck on a task:
    - “Do you see anything that might help you?”
    - “What do you think that button does?”
- Controls the session and prevents interruptions by observers
Observer’s role

- Be quiet!
- Don’t help, don’t explain, don’t point out mistakes
- Watch your body language
- Take notes
- Watch for critical incidents
  - Usually negative: errors, repeated attempts
- May be positive:
  - “Cool!”
  - “Oh, now I see.”
Critical incidents (1)
Critical incidents (2)
Recording observations

- Pen & paper notes
  - Prepared notes can help

- Audio recording

- Video recording
  - Usability labs usually contain 2 cameras (user’s face and screen)
  - User may be self-conscious
  - Good for closed-circuit view by observers in another room
  - Generates too much data
  - **Retrospective testing**: go back through the video with the user, discussing critical incidents

- Screen capture and event logging
  - Cheap and unobtrusive
  - Camtasia, CamStudio
How many users?

- Landauer-Nielsen model
  “A Mathematical Model of the Finding of Usability Problems”, INTERCHI’93
  - formula that relates the number of usability problems found by $i$ test users to the number of known problems $N$.

  \[ \text{Problems found} (i) = N(1-(1-\lambda)^i) \]

  - $\lambda$ is the probability of any user finding any problem.

- Which is better:
  - Using 15 users to find 99% of problems with one design iteration
  - Using 5 users to find 85% problems with each of 3 iterations
Flaws in Landauer-Nielsen model

○ Statistical considerations:
  ● different usability problems have different ‘visibilities’
  ● usability problems are often not independent

○ Consequences:
  ● \( \lambda \) may be much smaller than 31 %
  ● Spool & Schroeder study of a CD purchasing web site found \( L=8\% \), so 5 users only find 35% of the problems.
  ● You cannot predict with confidence how many users are needed to find most of the usability problems in your interface