Formative Evaluation of User Interaction: Before Evaluation Session

Topics

- Facilities and equipment
- Developing the experiment
- Developing the tasks
- Participant selection
- Preparing for participants

Facilities and equipment

- Usability lab
  - Formal lab is nice
  - One-way glass
  - Inter-com
  - Sound isolation
  - Low background noise (e.g., A/C)
  - Controlled lighting levels
  - Digital video and audio equipment, if necessary
  - Disability access
  - Emergency exits
Facilities and equipment

- Formal lab is nice
  - Reception room/area
    - To welcome and inform participants
    - For signing consent forms, etc.
    - For completing surveys and questionnaires

- Conference room works well
- Evaluator sits with participant
- You can still get a lot of useful usability data

- Observation area
  - Pipe video and audio into conference room
  - Better than one-way glass
  - Developers/clients separate from users
  - Gives developers team feeling
  - Afterward can sit and talk about it; can bring users in, too
  - A big way for developers to get the “usability religion”

Facilities and equipment

- Formal lab is not essential
- Conference room works well
- Evaluator sits with participant
- You can still get a lot of useful usability data

- Observation area
  - Pipe video and audio into conference room
  - Better than one-way glass
  - Developers/clients separate from users
  - Gives developers team feeling
  - Afterward can sit and talk about it; can bring users in, too
  - A big way for developers to get the “usability religion”

Facilities and equipment

- Data capture tools are very useful
- Scan converters are no longer useful now that we have direct screen capture programs like Camtasia.
- Pipe video/audio into conference room and have clients, developers observe (a real eye-opener)
Facilities and equipment

- **Practical hints**
  - High audio quality is essential & hard to get
    - No battery microphones – risk data loss if batteries go out during session
    - Need high quality (read: expensive) mike
    - Need separate amplifier/mixer
    - Need sensitive external mike, not built into camcorder
    - Place on participants lapel or on top of monitor

Developing the Experiment

- **Developing tasks**
  - Structured use: identification of representative, frequent and critical tasks
    - *Benchmark tasks*
      - Written out in detail, one per sheet, for participant
      - Usually take metrics during performance (for usability specifications)
      - You should already be started (for usability specifications)
      - See notes on usability specifications for details

- **Informal tasks: other tasks participant may perform, also written out in detail; no metrics are taken**
  - Exploratory use
    - No specific tasks
    - “Free play” for participant
  - All tasks, trials must be done from perspective of the user class being represented
Developing the Experiment

- Other details
  - Training materials, when appropriate
    - None should be needed for a calendar
    - Do up-front, be consistent
  - Props, task aids – have ready (e.g., telephone)

- Procedures
  - Usability lab and field evaluation (tradeoffs?)
    - Alpha and beta testing are NOT usability testing
  - Equipment
  - Typical length of time of evaluation session for one participant: 30 minutes to 4 hours, average 2 hours or less

- Participant selection
  - Representative users
    - Participants for each trial must match target user class for the associated benchmark task
    - Knowledgeable of target system domain
    - Know what they don’t like, but don’t usually know how to fix
Developing the Experiment

Participant selection

- User interaction design expert
  - Broadly knowledgeable in interaction development and use
  - Can find subtle problems
  - Can offer alternative suggestions for fixing problem

Developing the Experiment

Participant selection

- Someone old
  - Been around, knows how it’s done
- Someone new
  - From outside
- Someone borrowed
  - From a different department
- Someone blue
  - Never likes anything, always wants it different

Developing the Experiment

- How many participants is enough?
  - Focus not on large number of experiments with large number of users, but rather on extracting as much information as possible from every user
  - 1 participant is too few, more than 10 not worth it
  - Optimum number of participants is 3 to 5 per user class per major version/iteration (empirically-based rule-of-thumb)
Developing the Experiment

- Where does the '3 – 5 users' rule come from?
- Discovery likelihood = 1 - (1 - p)**n
  - Think of as % of problems found as function of # of users (n)
  - Each curve is for a given individual detection rate (p)

Developing the Experiment

- The '3 – 5 users' rule (1 - (1 - p)**n )

Developing the Experiment

- Then look at the expected number of new problems found for each added participant
- Cost is linear per added participant used
- Total # problems found / cost peaks out around 3-5 users
  - Based on lots of assumptions, such as average detection rate
  - Your mileage can vary
Developing the Experiment

- How many participants is enough?
  - More severe usability problems are typically detected by the first few participants
  - You could need many more participants
    - Sometimes you find numbers of new problems with 10th or 20th participant
    - Especially true for Web site testing (large scope; different users test different parts)

- Expected number of iterations per version is 3
  - Resource constraints often limit to less
  - Any iterations are better than none

- In subsequent cycles of evaluation, consider:
  - Keep “best” participant from previous cycle, add 2 (or more) new participants

- Preparing for participants
  - Develop instructions (see p.299-300 in book)
    - “You are helping us evaluate the system—we are not evaluating you!”
  - Develop informed consent form and non-disclosure agreements (see p. 306 in book, more about IRB soon)
Developing the Experiment

- Pilot testing
  - Pilot testing and rehearsal are essential
  - Design should not have known “show stopper” usability problems
- Establish evaluator roles
  - Facilitator – to keep evaluation session going
  - Observers – to help collect data
  - “Executor” – to run paper prototype

Developing the Experiment

- Bottom-line for developing the experiment:
  - Creativity rules
  - Ecological validity is important
  - These are just the basics; there are many variations on the theme
  - Do what you have to, to make it work (discover real usability problems)

The Institutional Review Board (IRB) *

- Purpose
  - Protect the rights of people participating in experiments
  - Protects university against liabilities (university is legally responsible for welfare of all human subjects involved in university activities)
The Institutional Review Board (IRB)

- Coverage
  - ALL empirical studies using human subjects (even usability testing you do for this class) conducted in the name of university (or any company) MUST be reported to and approved by the IRB
  - Most research in HCI does not put participants at risk and approval is given with ‘exempt’ status

The Institutional Review Board (IRB)

- Typically, principle investigator (team leader) must submit
  - “Request for exemption” cover letter
  - Statement of complete protocol
  - Written subject instructions
  - Informed consent form
  - Standard IRB forms
- I will submit a blanket application for the class
- You must remember to do this in the future

Informed Consent Form

- Legal requirements
  - Permission must be obtained prior to participation
  - Written document
  - Signed without duress or stress
  - In clear, understandable language
  - Copy given to participant
- Content
  - Statement of research purposes, procedures, duration of participation
Informed Consent Form

- Statement of ANY foreseeable risks or discomforts
- Statement of ANY benefits to participants (e.g., payment or education)
- Statement of alternative procedures
- Statement of confidentiality (anonymity of data)
- Full description of available treatments, if more than minimal risk

Signed consent forms must be retained for 3 years following IRB approval

Informed Consent Form

- Statement of persons to contact for answers to participant questions on research project, legal rights, and injuries
- Statement that participation is voluntary, state clearly any penalty (loss of benefits) if participant declines or discontinues participation (e.g., bonus fee for completing a whole series of tests)

Informed Consent Form

- IRB approval
  - Normally takes 1-2 weeks
  - May require changes in documents
  - Evaluates ethical and legal issues, not quality of the research

IRB approval
- Normally takes 1-2 weeks
- May require changes in documents
- Evaluates ethical and legal issues, not quality of the research