CS5714 Usability Engineering
Formative Evaluation of User Interaction: During Evaluation Session
Copyright © 2004 H. Rex Hartson and Deborah Hix

Topics

- Generating and collecting data
- Quantitative techniques
- Qualitative techniques
- Observational techniques
- Forms, variations, attitudes, tools

Generating and Collecting Data

- Preliminaries with participants
  - Explain protocol to participant, including any compensation
  - Show participant the lab and experimental set-up if they are interested
  - Have participant sign informed consent form and NDA
Quantitative Techniques

- Collecting quantitative data (to assess usability levels)
  - Benchmark tasks
    - Measuring time on task, number of errors, etc.
    - Quantitative measures such as timing can be valuable even with paper prototype, though not very precise
  - User satisfaction scores

Qualitative Techniques

- Collecting qualitative data (to identify usability problems)
  - Verbal protocol taking
    - Participants think aloud, talking while performing tasks
      - Can be intrusive, but effective
      - Some participants not good at talking
      - Evaluator/facilitator sits in room with participant to collect this kind of data
    - Can be used for both time and un-timed tasks
      - Studies show it can be done with minimal effect on performance time
  - Facilitator may need to prod participant who stops talking, but don’t get into discussion during timed tasks
  - Answer questions about what to do with a hint, not a direct answer
Qualitative Techniques

- Collecting qualitative data (to identify usability problems)
  - Critical incident taking
    - Critical incident: something that happens while participant is working that has significant effect on task performance or user satisfaction
    - Critical incidents are indicators of usability problems
    - Later analyze the problem and cause within the interaction design

Arguably single most important kind of formative evaluation data
Although participant may indicate a critical incident, it is responsibility of evaluator to identify and record critical incidents
In most situations, finding usability problems is bad, but in formative evaluation, finding usability problems is good

Pay attention to detailed participant behavior – IMPORTANT!
- It's easy to miss them! It's a skill; takes experience
- Example: user wasn't sure what the alarm clock icon meant
  - Could have had to do with time of day. Solution: show it "ringing" to emphasize alarm part
- Example: user confused by word “cancel” on button in dialogue box showing appointments
  - Subtle: Normally understand “Cancel”, but with appointment in calendar domain “cancel” has meaning of removing appointment
Observational Techniques

- Some observational data collection techniques
  - Structured interviews
    - Post-session questioning
    - Typically obtain general information
  - Co-discovery
    - More than one participant, using system together, thinking aloud together
    - Can lead to rich verbal protocol from conversations among participants

Observational Techniques

- Some observational data collection techniques
  - Software tools for critical incident recording
  - Note taking – the primary technique
    - Most important: Real-time notes (e.g., pencil and paper, online)
    - Nothing beats this for effective data gathering

Observational Techniques

- Some observational data collection techniques
  - Audio recording can be useful
    - Effective if used selectively for note taking, if not too distracting
    - Can be used to capture continuous dialogue with participant (more agile than video taping)
Observational Techniques

- Some observational data collection techniques
  - Video taping
    - Used primarily as backup
    - Captures every detail, but tedious to analyze
    - Generally one camera on hands/keyboard/mouse/screen;
      if a second camera, on user's face
    - Used to use scan converter for screen action (but
      resolution too low to read)

- The role of video taping today
  - Screen capture software gives much higher
    resolution screen images (e.g., Camtasia)
  - Typically not use video camera (unless important
    – e.g., VR)
  - Video can be used with paper prototypes, too

Data Collection Forms

- Form for collecting both quantitative and
  qualitative data during session

DATA COLLECTION FORM

<table>
<thead>
<tr>
<th>Task Name:</th>
<th>Participant ID:</th>
<th>Task Start Time:</th>
<th>Task End Time:</th>
<th>Task Duration:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tape
Counter
Critical Incident Description
Adopt the Right Attitude

- Evolution of developers’ attitude as they watch user evaluating a product
  
  “Stupid user!”
  “Let me at him/her!”
  “It’s his/her (another developer’s) fault!”
  “I’m mud!”
  “Let’s fix it!”

Variations

- Variations on the theme
  
  - Major point: No rules; do what works best in your situation
  - Evaluator sitting with participant (cf. in separate room)
  - Abandon verbal protocol if it doesn’t work for a participant
  - Try co-discovery with two participants

- **Video-clip**: Formative evaluation of Envision

Data Collection Tools

- Ideal – a software tool for critical incident recording
  
  - Usability engineer uses to capture raw usability data
  - Quantitative: timing, error counts
  - Critical incidents: most important
    - Creates database record of raw usability data
    - Tags record with video (Camtasia) clip for later review
    - Gathers critical incident records for session in database
    - Shares database with other UE tools
    - Feeds critical incident descriptions to usability problem analysis tools
Team Exercise – Formative Usability Evaluation

- Goal: To perform the data collection part of a very simple formative usability evaluation
- Activities:
  - Assemble in teams

Team Exercise – Formative Usability Evaluation

- Decide roles for team members:
  - Prototype executor, to move transparencies, provide feedback (person who knows design best and who can "play computer")
  - Evaluation facilitator, to keep experiment moving, to interact with participants, and to record critical incidents (qualitative data)
  - User performance timer, to time participants performing tasks and/or count errors (quantitative data)
  - Two participants, to trade to another team
  - Anyone else can help record critical incidents

Team Exercise – Formative Usability Evaluation

- Now make the switch:
  - Trade your two participants to another team, getting two new participants from a different team (we'll help make this work in a "cycle" among the teams)
  - Your new participants are now permanently on your team (for this exercise)
  - Newly formed teams sit together in groups now
  - We will now pass out and explain the forms you will use
Team Exercise – Formative Usability Evaluation

Getting ready:
- Have your new participants leave the room temporarily
- Get your prototype “booted up”
- Bring first participant into “lab”, greet them, and explain evaluation session to them

Cautions and restrictions
- Team members must not coach participants as they perform tasks.
- Person playing computer must not anticipate user actions, especially do not give the correct computer response for a wrong user action! Respond only to what user actually does!
- Person playing computer may not speak, make gestures, etc.
- You may not change the design on the fly!

How much data to collect?
- You need to collect about a dozen or more critical incidents in this exercise
- If you do not get at least a half dozen from each user, continue with that user doing exploratory use of your prototype until you get enough critical incidents
- Have them browse through each screen, looking at each object (button, menu, etc.) commenting on and giving their opinion about usability of various features
Team Exercise – Formative Usability Evaluation

Run the experiment:
- Have first participant use your prototype to perform the benchmark tasks for your objective usability specifications. That is, have participant read first task aloud, then perform that task while thinking aloud.
  - As executor moves transparencies in response to participant actions, the timer records times and/or counts errors as user performs task. Evaluation facilitator records critical incidents.
  - Don’t count reading of task in task timing.

- Next have participant read second task aloud and perform it while thinking aloud.
- Have this participant complete questionnaire, and then give them their “reward”.
- Now have the second participant perform the tasks and complete the questionnaire. The first participant should stay and help with observations.

Deliverables: none
Schedule: Complete by end of class