Lecture 5

Usability Evaluation Methods (other than Usability Testing) (chapter 4)

Usability Evaluation Summary

• Purpose: Evaluation for usability
• Methods
  – Without Users (analytic)
    • Guidelines (Chapter 2.2)
    • Interface Walkthrough
    • Expert Review (Chapter 4.2)
    • Model-Based analysis (Keystroke Model)
  – With Users (empirical)
    • Usability testing (Chapter 4.3)
    • Experiments (Chapter 4.7)
    • Field Studies (Chapter 4.5)
    • Surveys (Chapter 4.4)

Testing Goals vs. Method

<table>
<thead>
<tr>
<th>Testing Goals</th>
<th>Guidelines</th>
<th>Walkthrough</th>
<th>Expert</th>
<th>Keystroke Model</th>
<th>Usability Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completeness</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correctness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Consistency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Interface (Cognitive) Walkthrough

- **Definition**
  - Given an interface and a set of functions, for each function the developers walk through the user’s action sequence noting any problems.

- **Purpose**
  - Checks usability by a detailed rehearsal of steps in user tasks

- **Problems**
  - Incomplete since the developers may have made oversights
  - Developers are not the users

Expert (Heuristic) Evaluation

- **Definition**
  - Human factors expert evaluates interface
  - Uses Guidelines and experience
  - Suggests changes

- **Purpose**
  - Use human expertise to find usability problems

- **Problems**
  - Expert may be unavailable
  - Experts may disagree
  - An expert is not a user

Heuristic evaluation
(Niesen & Molich, CHI 1990, “Heuristic evaluation of user interfaces”)

- **Problem**
  - Individual evaluators found between 20-51% of usability problems

- **Solution**
  - Use aggregated results from 3 to 5+ evaluators
Keystroke Level Model

• Definition
  – Predicts time to do a task for an expert user
• How to do it
  – Specify a task with low-level actions
  – Give times for each action and system response
    • key press = .2 sec; mouse point = 1.1 sec; reach = .4 sec
  – Sum to compute estimated time for the task
• Problems
  – Reliability of prediction is 80%
  – Cannot predict learning time

Guidelines

• Shared language
• Best practices
• Proponents
  – Encapsulates experience of real designers

Guidelines

• Definition
  – Guidelines are written “standards” and heuristics (rules of thumb) for interfaces.
    Following them should lead to more usable designs.
• Guidelines are used to evaluate the mockups, scenarios, storyboards and user interaction networks specifications
Guidelines (Apple Computer)

• Examples from Apple Computer’s *Human Interface Guidelines*, 1985-1989

  – Learnability. Use concrete metaphors.

Guidelines of Usability Experts
(Arnold Lund, Ameritech, 1995)

• Experts in HCI design field suggested rules of thumb they found particularly useful during design
• 34 rules defined
• 31 HCI experts rated each of the rules of thumb by their estimate of magnitude of impact on usability of designs
• 5 is most impact; 1 is least
• Forced choice (20% of rules “5”, 20% “4”, etc.)
• 82% inter-rater correlation: high consensus

Guidelines of Usability Experts
(5 is most impact on usability)

4.1 Know thy user, and YOU are not the user.
4. Things that look the same should act the same.
3.1 The information for the decision needs to be there when the decision is needed.
3.8 Error messages should actually mean something to the user, and tell the user how to fix the problem.
3.4 Every action should have a reaction.
3.7 Don’t overload the user’s buffers.
3.6 Consistency, consistency, consistency.
3.5 Minimize the need for a mighty memory.
3.4 Keep it simple.
Guidelines of Usability Experts - cont.

3.4  The more you do something, the easier it should be to do.

3.4  The user should always know what is happening.

3.4  The user should control the system. The system shouldn't control the user. The user is the boss, and the system should show it.

3.1  The idea is to empower the user, not speed up the system.

3.3  Eliminate unnecessary decisions, and illuminate the rest.

3.3  If I made an error, let me know about it before I get into real trouble.

3.3  The best journey is the one with the fewest steps. Shorten the distance between the user and their goal.

Guidelines

• Problems
  – Too specific, incomplete, hard to apply
  – Unsystematic folklore: sometimes wrong
  – Tradeoffs

Dilbert
by Scott Adams
Testing & Evaluation: What is the best method?  
(Jeffries et al., CHI ’91)

- Which method discovers the most severe usability problems?
  - Walkthrough: 3.44
  - Expert Evaluation: 3.59
  - Guidelines: 3.61
  - Usability Testing: 4.15

- Range: 1= trivial  9= critical
- mean shown for severity of problems discovered

<table>
<thead>
<tr>
<th>Method</th>
<th>Most Severe</th>
<th>Least Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walkthrough</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Guidelines</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Usability Testing</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>Expert Evaluation</td>
<td>28</td>
<td>52</td>
</tr>
</tbody>
</table>

- most= top third  least= bottom third

Testing & Evaluation: What is the best method?

- Number of severe problems found

<table>
<thead>
<tr>
<th>Method</th>
<th>Most Severe</th>
<th>Least Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walkthrough</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Guidelines</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Usability Testing</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>Expert Evaluation</td>
<td>28</td>
<td>52</td>
</tr>
</tbody>
</table>

Usability Testing in Practice  
(Nielsen, *Usability Engineering*, 1993)

Survey of 28 projects from broad variety of companies and industries

Findings:

<table>
<thead>
<tr>
<th>Metric</th>
<th>Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project size in person-years</td>
<td>24</td>
</tr>
<tr>
<td>Actual share of budget for usability</td>
<td>6%</td>
</tr>
<tr>
<td>Ideal share of budget for usability</td>
<td>10%</td>
</tr>
<tr>
<td>Actual usability effort in person-years</td>
<td>1.5</td>
</tr>
<tr>
<td>Ideal usability effort in person-years</td>
<td>2.1</td>
</tr>
</tbody>
</table>
Usability Testing in Practice cont.

• Findings:
  – Usability effort is independent of project size!
  – Why? Many usability activities take about the same time to perform, no matter how difficult the program is to develop

Testing and Evaluation of Users Summary

• Methods
  – Without Users
    • Guidelines
    • Interface (Cognitive) walkthrough
    • Heuristic evaluation
    • Keystroke Level Model
  – With Users
    • Usability testing
    • Questionnaire & Interview
• Use all methods for a more usable interface

Beyond Guidelines: Principles

• More fundamental, widely applicable, and enduring than guidelines
• Need more clarification
• Fundamental principles
  – Principle 1: Determine user’s skill levels
  – Principle 2: Identify the tasks
  – Principle 3: Five primary interaction styles
  – Principle 4: Eight golden rules of interface design
  – Principle 5: Prevent errors
  – Principle 6: Automation and human control
Principle 1: Determine user’s skill levels

- “Know thy user” Hansen (1971)
- Age, gender, physical and cognitive abilities, education, cultural or ethnic background, training, motivation, goals and personality
- Design goals based on skill level
  - Novice or first-time users
  - Knowledgeable intermittent users
  - Expert frequent users
- Multi-layer designs

Principle 2: Identify the tasks

- Task Analysis usually involves long hours observing and interviewing users
- Decomposition of high level tasks
- Relative task frequencies

<table>
<thead>
<tr>
<th>Task</th>
<th>Occurrence</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task A</td>
<td>0.10</td>
<td>1.0</td>
</tr>
<tr>
<td>Task B</td>
<td>0.50</td>
<td>2.0</td>
</tr>
<tr>
<td>Task C</td>
<td>0.30</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Principle 3: Choose an interaction style

- Direct Manipulation
- Menu selection
- Form fillin
- Command language
- Natural language
Principle 4: The 8 golden rules of interface design

1. Strive for consistency
2. Cater to universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions
7. Support internal locus of control
8. Reduce short term memory

Principle 5: Prevent errors

- Make error messages specific, positive in tone, and constructive
- Mistakes and slips (Norman, 1983)
- Correct actions
  - Gray out inappropriate actions
  - Selection rather than freestyle typing
  - Automatic completion
- Complete sequences
  - Single abstract commands
  - Macros and subroutines
Principle 6: Automation and human control

Automation and human control (cont.)

- Successful integration:
  - Users can avoid:
    - Routine, tedious, and error prone tasks
  - Users can concentrate on:
    - Making critical decisions, coping with unexpected situations, and planning future actions

Beyond Principles: Theories

- Beyond the specifics of guidelines
- Principles are used to develop theories
- Descriptions: explanatory or predictive
  - Example: Fitts Law for predicting pointing time
- Motor task, perceptual, or cognitive