Content today

- Method Assessment
- Rationale for user involvement
- Different design approaches with different forms of user involvement
- Different techniques
- Choosing and combining methods
Why Do We Evaluate In HCI? (1)

1. Evaluation to produce generalized knowledge
   - are there general design principles?
   - are there theories of human behaviour?
     - Explanatory
     - Predictive
   - can we validate ideas / visions / hypotheses?

Why Do We Evaluate In HCI? (2)

2. Evaluation as part of the Design Process
How Do We Evaluate In HCI? (1)

A. Pre-design
- what do people do?
- how can we understand what we need in system functionality?

evaluation produces
- key tasks and required functionality
- work practices
- organizational practices
- user type...

How Do We Evaluate In HCI? (2)

B. During initial stage, developing design ideas and representations
- evaluate choices of initial design ideas and representations
  - is the representation appropriate?
  - does it reflect how people think of their task

evaluation produces:
- user reaction to design
- validation and list of problem areas (bugs)
- new design ideas
How Do We Evaluate In HCI? (3)

C. During iterative development, refining a design / representation
   - fine tune the interface, looking for usability bugs
     - can people use this system?

   Evaluation produces:
   - user reaction to design
   - validation and list of problem areas (bugs)
   - variations in design ideas

How Do We Evaluate In HCI? (4)

D. Post-design
   - acceptance test: did we deliver what we said we would?
     - verify that human/computer system meets expected performance criteria
     - ease of learning, usability, user’s attitude, performance criteria
     - e.g., a first time user will take 1-3 minutes to learn how to withdraw $50. from the automatic teller

   - revisions: what do we need to change?

   - effects: What did we change in the way people do their tasks?

   Evaluation produces
   - testable usability metrics
   - actual reactions
   - validation and list of problem areas (bugs)
   - changes in original work practices/requirements
Evaluation AND Design in HCI!

- **Design and evaluation**
  - Best if they are done **together**
    - evaluation suggests design
    - design suggests evaluation
    - use evaluation to create as well as critique
  - Design and evaluation methods **must fit** development constraints
    - budget, resources, time, product cost...
    - do triage: what is most important given the constraints?
  - Design usually needs quick approximate answers
    - precise results rarely needed
    - close enough, good enough, informed guesses,…

**User Population**

- ‘All’? or a few?
User Representation

- Expert?
- Or real user?

Timing of Involvement

- Early?
- And/or late?
Location of Analysis

- Laboratory?
- In context?

Design Approaches

Review:
- User-Centred Design (UCD)
- Participatory design (PD)
- Socio-Technical Design (STD)
- Soft-Systems Methodology (SSM)
- Joint Application Design (JAD)

=> differences in cultures, underlying theories, different countries.
Similarities and Differences

Background:
♦ Time/origin (when/ country)
♦ Rationale for involvement
♦ Background theory

Attributes:
♦ User representation
♦ User role (how)
♦ User control (how)
♦ User Involvement timing (when)

Non-configurable Attributes

<table>
<thead>
<tr>
<th>Attributes</th>
<th>UCD</th>
<th>PD</th>
<th>STD</th>
<th>SSM</th>
<th>JAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Late 1970s</td>
<td>Late 1960 - early 1970s</td>
<td>Early 1970s</td>
<td>1970s</td>
<td>Late 1970s</td>
</tr>
<tr>
<td>Rationale</td>
<td>Pragmatic</td>
<td>Political, Philosophical</td>
<td>Political, Commitment and buy-in</td>
<td>Pragmatic (Political)</td>
<td>Pragmatic</td>
</tr>
<tr>
<td>Theory</td>
<td>Human information processing</td>
<td>Labour relations, group learning</td>
<td>Socio-technical systems theory</td>
<td>Systems thinking</td>
<td>Group dynamics, software engineering</td>
</tr>
</tbody>
</table>
Configurable Attributes

<table>
<thead>
<tr>
<th>Attributes</th>
<th>UCD</th>
<th>PD</th>
<th>STD</th>
<th>SSM</th>
<th>JAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User represent.</td>
<td>Repres-</td>
<td>All future</td>
<td>All future</td>
<td>Repres-</td>
<td>Repres-</td>
</tr>
<tr>
<td></td>
<td>entatives</td>
<td>users</td>
<td>users</td>
<td>entatives</td>
<td>entatives</td>
</tr>
<tr>
<td>HOW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User represent.</td>
<td>Indirect</td>
<td>Direct</td>
<td>Direct</td>
<td>Indirect,</td>
<td>Direct</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>direct</td>
<td></td>
</tr>
<tr>
<td>User influence</td>
<td>Product</td>
<td>Process,</td>
<td>Process,</td>
<td>Product</td>
<td>Product</td>
</tr>
<tr>
<td></td>
<td></td>
<td>product</td>
<td>product</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHEN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User involv.</td>
<td>Mostly early and late</td>
<td>Through-out</td>
<td>Through-out</td>
<td>Mostly early</td>
<td>Mostly early</td>
</tr>
<tr>
<td>timing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What is a Method? (1)

- Method definition:
  - Formalized procedures / tools that guide and structure the process of gathering and analyzing information.
  - Different methods can do different things.
  - Each method offers potential opportunities not available by other means.
  - Each method has inherent limitations...
What is a Method? (2)

All methods:
- enable but also limit what can be gathered and analyzed
- are valuable in certain situations, but weak in others
- have inherent weaknesses and limitations
- can be used to complement each other’s strengths and weaknesses.

-McGrath (Methodology Matters, 1995)

Why Use Different Methods?

- Information requirements differ
  - Pre/post-, iterative design, generalizable knowledge...
- Information produced differs
  - outputs should match the particular problem/needs
- Cost/benefit of using method
  - cost of method should match the benefit gained
- One method’s strength can complement another’s weakness
  - no one method can address all situations
- Constraints
  - may force you to choose quick and dirty discount usability methods
How Can We Compare Methods? (1)

- **Relevance**
  - does the method provide information to our question / problem?

- **Naturalistic**
  - is the method applied in an ecologically valid situation?
    - observations reflect real world settings: real environment, real tasks, real people, real motivation

- **Generalization**
  - how well can I generalize the information produced to other situations?

- **Repeatability**
  - would the same results be achieved if the test were repeated?

How Can We Compare Methods? (2)

- **Validity**
  - External validity: can the results be applied to other situations?
  - Internal validity: do we have confidence in our explanation?

  *Does the test measure something of relevance to usability of real products in real use outside of lab?*

  - Some typical reliability problems of testing vs real use
    - non-typical users tested
    - tasks are not typical tasks
    - physical environment different
      - quiet lab vs very noisy open offices vs interruptions
    - social influences different
      - motivation towards experimenter vs motivation towards boss
How Can We Compare Methods? (3)

- **Quickness**
  - can I do a good job with this method within my time constraints?
- **Cost**
  - Is the cost of using this method reasonable for my question?
- **Equipment**
  - What special equipment / resources required?
- **Personnel, training and expertise**
  - What people / expertise are required to run this method?

How Can We Compare Methods? (4)

- **Subject selection**
  - how many do I need, who are they, and can I get them?
- **Scope of subjects**
  - is it good for analyzing individuals? small groups? organizations?
- **Type of information** (qualitative vs quantitative)
  - is the information quantitative and amenable to statistical analysis?
- **Comparative**
  - can I use it to compare different things?
- **Routine application**
  - is there a fairly standard way to apply the method to many situations
How Can We Compare Methods? (5)

- **Measures**
  - can I see processes or outcomes
- **Metrics**
  - are there useful, observable phenomena that can be measured
- **Control**
  - can I control for certain factors to see what effects they have?
- **Organizational**
  - can they be included within an organization as part of a software development process
- **Politics**
  - are there ‘method religion wars’ that bias method selection?

Choosing Methods

- **Considerations:**
  - Available time
  - Expertise
  - Access to users
  - Type of research question
  - Design problem / product
  - Design project phase (representations available)
ISO/TR 16982: Comparing methods

- Characteristics being considered:
  - user (access to users, workplace, handicaps)
  - task (complexity, error critical, new)
  - product (simple/complex, diverse uses)
  - project (time, costs, early diagnosis)
  - skills (human factors skill [not] available)
Example-1: method selection (1)

- Industrial scraps treatment software
- Small company
- Software about own competences
- To be used by novices and experts (within company)
- Emphasis on usability
- Duration: 5 days/man

Example-1: method selection (2)

Table B.3: Grid of the ISO 16882: Industrial scraps treatment software

© M. Rauterberg, 2006
JFS-USI Primer-10
29/40

© M. Rauterberg, 2006
JFS-USI Primer-10
30/40
Example-2: method selection (1)

- Web site documentation centres and libraries
- Information about centres and catalogue
- Two user groups:
  - Professionals of documentation
  - Users of centres and libraries
- Goal: improve site
- Duration: 60 days/man

Example-2: method selection (2)
Combining methods

- Objective and subjective measures:
  - e.g. observations and questionnaires
  - audio-set study
- Without and with users
  - e.g. expert review and usability test
  - KPN case
- Meet requirements and new design ideas
  - e.g. usability test and post-walkthrough
  - USI example

Case-1: Personal internet page (1)

Test acceptance and usability of the personal internet page (by KPN research)
- target group are inexperienced and experienced internet users
- prototype available on the internet, IE 4 needed
- confidentiality very important
- time 400 hour, 2 persons
- start half June, prototype ready begin July, results must be ready 1 September
Case-1: Personal internet page (2)

Case-1: Personal internet page (3)
Approach for Case-1?

- Which test methods would you apply?
- Would you invite test users, and if so how many?
- Who would you invite?
- What would be your global project planning?

Case-1: Personal internet page (4)

Usability test in lab with people from KPN, middle July
- 14 people in U-lab at Research (minimum 12)
- 5 inexperienced, 4 little experience, 5 internet users
- walk through scenario’s with questions
- first impression and general impression after use

Acceptance pilot, people from KPN, July - August
- 12 people, 8 internet users
- 3 weeks use of the page at home
- questionnaire first impression, one hour use and general impression after 3 weeks
Summary

- Different forms of user involvement
- Choosing and combining methods
  - With and without users
  - Subjective and objective
  - Evaluation and design ideas

References

ISO/TR 16982, *Ergonomics of human-system interaction – Usability methods supporting human-centered design*

UCD at IBM:

UCD Works (2002):
http://www.idemployee.id.tue.nl/g.w.m.rauterberg/presentations/2003%20UCDworks_files/frame.htm
http://www.idemployee.id.tue.nl/g.w.m.rauterberg/videos.html#C

Usability Net:
http://www.hostserver150.com/usability/home.htm