Designing Interactive Systems

rooms: B109 – B107
time: 14h00 – 17h00

Wendy E. Mackay
Tuesday, 8 January 2012
lecture 5

Today

Homework: Midterm assessment (group)  Video prototype (group)

Lectures: Evaluation: Qualitative and Quantitative

Class work: Structured Walkthrough
Experiment design

due Jan. 15: Revise your storyboard:
Add three interaction points
Identify two potential breakdowns
Identify three improvements

due: January 8, 2013

Mid-term assessment:
Section I: Who is your user?
2 pages
user profile, key incidents from interviews
user personas, use scenario

Section II: What are your design inspirations?
½ page
brainstormed ideas, web search

Section III: What is your design concept?
2 pages
design concept, alternatives
function-interaction table, design scenario

Section IV: Storyboard
2 A3 pages

Section V: Video prototype
5-7 minutes

Contacts

Wendy E. Mackay
Research Director, in|situ| lab, INRIA
mackay@in.fr

Jérémie Garcia
Ph.D. candidate, in|situ| lab, INRIA
garcia@in.fr

Go to: insitu.lri.fr/People/Mackay
and click on: Design and Evaluation of Interactive Systems
or
Website: http://insitu.lri.fr/People/DesignAndEvaluationOfInteractiveSystems
Generative Design

<table>
<thead>
<tr>
<th>Discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who is the user?</td>
</tr>
<tr>
<td>Invention</td>
</tr>
<tr>
<td>What is possible?</td>
</tr>
<tr>
<td>Design</td>
</tr>
<tr>
<td>What should it be?</td>
</tr>
<tr>
<td>Evaluation</td>
</tr>
<tr>
<td>Does it work?</td>
</tr>
<tr>
<td>Redesign</td>
</tr>
<tr>
<td>How to improve it?</td>
</tr>
</tbody>
</table>

Evaluation: Does it work?

- Collect information
  - Design Walkthrough
  - Experiment
- Analyse information
  - Qualitative
  - Quantitative
- Resources for design
  - List of problems, issues
  - Implications for redesign

Evaluation Techniques

- You can evaluate anything … qualitative and quantitative
- Our focus is on evaluating prototypes but we'll touch on techniques for more finished systems
- Design walkthroughs
  - Heuristics
  - Design principles
- User studies
  - Wizard of Oz studies
  - Controlled experiments
  - Field studies
Evaluation strategies

- Introspection
  Designers test prototypes themselves

- Theoretical analysis:
  GOMS, CIS, QOC, Claims analysis, Task analysis

- Heuristic (rule-based) evaluation
  Designers apply rules

- Usability studies with users
  Lab or field studies

- Design walkthroughs
  Systematic qualitative evaluation

Introspection

Designer’s test of his or her own system
Contrast this to introspection as a design technique
- not what it should be
- rather what is wrong
Least likely to identify key user problems

Some advice:
- Be systematic: perform specified tasks, follow a scenario
- Videotape or take notes as you go
- otherwise you will forget the details
- Highlight problems but also new ideas

Theoretical analysis

Calculate each interaction

Heuristic (rule-based) evaluation

Common in industry
because it is inexpensive
Limited in scope
does not handle rules that conflict
Often too vague:
better at identifying problems
than helping to create solutions
Nielsen’s argument: a few users are ok

How many tests are necessary?

<table>
<thead>
<tr>
<th>Number of tests</th>
<th>Easy to detect</th>
<th>Average to detect</th>
<th>Hard to detect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>1</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>2</td>
<td>0.7</td>
<td>1.0</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Highly cited, but largely discredited by researchers!

Usability Study

Observe users using technology in a real-world setting
Focus is on finding out about the system (contrast with phase I observation where focus on the user)
Of course, it is really about the user’s interaction with the system
Note: You can evaluate partial prototypes before creating full working systems

Working with real users...

is often difficult
How do you define a ‘step’
What if the user can’t perform the task
How important is previous experience?
What about learning effects?
What if it works differently in different contexts?
Even so, it is almost always worth the trouble

Lab tests

Ask users to try the system
Follow a scenario
Execute particular tasks
Perform specified functions

Capture their actions:
video or log

Afterwards:
Debrief them
(explain the study)
Interview them
(find out their opinions)
**Field test**

Deploy the system with real users, in the context of their everyday lives, for days, weeks or months.

Example: \( \text{WAVE link between England and the Netherlands} \)

<table>
<thead>
<tr>
<th>Logging data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record every keystroke while user performs specified tasks</td>
</tr>
<tr>
<td>Careful: low-level logs are hard to interpret</td>
</tr>
<tr>
<td>Create high-level logs where possible</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Design Walkthrough</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step-by-step evaluation of sequential material to identify as many problems as possible at each step</td>
</tr>
<tr>
<td>Similar to brainstorming: Goal is to identify maximum quantity of problems</td>
</tr>
<tr>
<td>Contrast with brainstorming: Do not defer judgement</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Design Walkthrough</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of comments:</td>
</tr>
<tr>
<td>Focus on material, not author</td>
</tr>
<tr>
<td>Constructive not destructive</td>
</tr>
<tr>
<td>Specific, not general</td>
</tr>
<tr>
<td>Problems then questions then suggestions</td>
</tr>
<tr>
<td>Examples:</td>
</tr>
<tr>
<td>“The text is too small to read”</td>
</tr>
<tr>
<td>“The user can’t see where to change the setting”</td>
</tr>
<tr>
<td>“That task takes four steps”</td>
</tr>
<tr>
<td>Authors: Accept the problems, but do not discuss solutions! Try to <strong>find</strong> as many issues as possible – don’t solve them.</td>
</tr>
</tbody>
</table>
### Design Walkthrough

**Appropriate for many types of material**

- Originally for programmers and their code

**However it works well for:**
  - Text documents: articles, manuals, specifications, reports
  - Design resources: design scenarios, storyboards, paper prototypes, video prototypes

### Group characteristics:
- **peers** should do other types of evaluations
- **small** 4-8 works well
- **diverse** include diverse perspectives

In addition to your personal opinion adopt specific roles:
- **technical** Is there an error or problem?
- **user** Is it hard to do?
- **manager** Is this function necessary?

**or apply a set of design rules, principles or perspectives:**
- Norman's rules
- Shneiderman's rules

(We'll cover others later)

### 3 principles from Norman (1990)

Recommendations for creating 'good' conceptual models

1. **Visibility principle:** Make things visible
   - Help users understand the state of the system by observing the interface
2. **Mapping principle:** Link actions and effects
   - Help users understand the correspondence between actions and results
   - Controls and their effects
   - The state of the system and what is visible
3. **Feedback principle:** Inform the user
   - Feedforward: what is possible
   - Feedback: what just happened

### 8 design rules from Shneiderman (1998)

1. **Consistency**
2. **Short-cuts for expert users**
3. **Informative feedback**
4. **Design dialogues with closures**
5. **Prevent errors and help repair them**
6. **Provide reversible operations**
7. **Give control to the user**
8. **Reduce short-term cognitive load**
### Exercise: Design Walkthrough

**Goal**
Rapid identification of potential problems, based on different perspectives and design criteria

**Roles**
- **Author**: presents / chooses type of critique
- **Participants**: find as many problems as possible

**Procedure**
- Small group of peers with different roles & expertise
- 15-60 minute sessions
- Everyone gets a copy
- Group identifies as many problems as possible
- Use roles or rules to find more issues

### When to use experiments?

Experiments are good for helping to answer **focused** questions
- Compare at least two things

Careful:
- usability study ≠ experiment

### Simple experiment

**Goal**
Choose the best design alternatives by watching users try the prototype

**Procedure**
- Describe the design objective
- Identify several alternatives
- Choose the independent and dependent
- Specify the null hypothesis and make a prediction
- Set up the test conditions to compare each condition
- Use at least three real users
- Analyse the results: are they significantly different?

### Design a simple experiment

1. Specify the **functionality** offered to users
   - What does the system do?
2. Specify **alternative interaction techniques**
   - How does the user accomplish it?
3. Specify the **independent variables** (factors)
   - Experimenter decides on the values
4. Specify the **dependent variables** (measures)
   - User behavior determines the values
5. Specify the tasks the user will perform (**operationalize behavior**)
   - Specify experimental and control groups
6. Specify appropriate statistical tests
   - Is the difference real?
Design a simple experiment

- Specify the **functionality** offered to users
  - What does the system do?

- Specify **alternative interaction techniques**
  - How does the user accomplish it?

- Specify the **independent variables** (factors)
  - Experimenter decides on the values

- Specify the **dependent variables** (measures)
  - User behavior determines the values
### Design a simple experiment

1. **Specify the functionality offered to users**
   - What does the system do?

2. **Specify alternative interaction techniques**
   - How does the user accomplish it?

3. **Specify independent variables (factors)**
   - Experimenter decides on the values

4. **Specify dependent variables (measures)**
   - User behavior determines the values

5. **Specify the tasks the user will perform (operationalize behavior)**
   - Specify experimental and control groups

---

### Example

**Comparison of linear and circular menus**

**Null hypothesis (H₀)**

There is no difference in performance between users in terms of time or error when selecting an item from a linear or a circular menu regardless of type of menu, number of menu items or previous experience.
Specify the independent variables

<table>
<thead>
<tr>
<th>Independent variables (factors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>are those we want to vary or control</td>
</tr>
<tr>
<td>The combinations of variables define the <strong>conditions</strong></td>
</tr>
</tbody>
</table>

**Independent variables:**
- **Type of menu:** linear, circular
- **Number of items:** 3, 6, 9, 12, 15
- **Expertise:** expert, novice, intermittent

**Values:**

\[2 \times 5 \times 3 = 30 \text{ unique conditions}\]

**Caution:** Most user-related variables are dependent on the user's behavior. User experience is an independent variable, chosen independently by the experimenter.

Specify the dependent variables

**Dependent variables depend on the user’s behavior**

Also called measures because they measure user’s behavior

For a reliable statistical test, you need sufficient measures per condition.

- **Rule of thumb:**
  - ~12 for small-n statistics (student’s t)
  - ~30 for normally distributed tests

**Dependent variable might include:**
- Time to select an item
- Number of errors
- What else?

Operationalize the behavior

Trickiest part of the experimental design but when it’s done well, it seems obvious

Simplify the task as much as possible, to eliminate bias and external factors, without making it unrealistic.

Example: Fitts’ pointing task
- only one dimension
  (target = vertical band)
- reciprocal pointing
  (back and forth between 2 targets)

Controlled experiment

Which gesture is faster, with fewer errors, for the user?
- “EyeToy” Move your hand to indicate a choice
- Object Tracker Place a colored object over the choic
**Operationalize the behavior**

Define the specific menu selection task
Ensure that the conditions are as similar as possible:
- Same labels for the menu items
- Same location of the menu on the screen (centered)
- Highlight the item to select (to avoid searching for it)

**Run the experiment**

Prediction:
- Always write your subjective predictions before you discover the results
- Another example of looking for surprises

Control any factors that might bias the results:
- All subjects receive the same instructions
- All subjects perform tasks under the same conditions
- All instructions are simple and clear
- Informal contact kept to a minimum

Double blind experiment:
- Neither the experimenter nor the subject know which group receives which treatment
- Can our experiment be double blind?

**Run the experiment**

Obtain informed consent from the subjects
Ensure that subjects remain anonymous
Associate a number with each subject
Choose conditions based on those numbers
Gather experimental data
- Test that they are reliable and valid
- Minimize data treatment and preserve raw data

**Prediction ≠ Null Hypothesis**

For this experiment:
- I think that circular menus will be faster than linear menus regardless of experience and the number of menu items
- Linear menu performance will decrease with more items
- Circular menu performance will drop as more items are added
Collect data

Ensure that the data log is human-readable yet easy to analyze by both people and machines.

Start S1 E C-L 3-12-15-9-6 Mon 21 Nov 2012 15:45:54  
Condition S1 E C 3 Mon 21 Nov 2012 15:46:35  
# sujet expertise type taille item hit/miss tps(ms)  
Trial S1 E C 3 2 Hit 1254  
Trial S1 E C 3 1 Miss 885  
...  
Condition S1 E C 12 Mon 21 Nov 2012 15:54:22  
Trial  
...  
End S1 E C-L 3-12-15-9-6 Mon 21 Nov 2012 16:23:55

Cinematic log

CPN 2000  
Experiment

Analysis of Variance (ANOVA)

Compare:  
Marking menus  
Floating palettes  
Tool glasses

Tracking performance

Compare:  
Orthozoom and SDAZ  
Movement time (MT) by Index of Difficulty (ID)
Summary

Controlled experiments
- Manipulate independent variables
- Observe the effect on dependent variables
- Draw conclusions from statistical tests
  (but also qualitative responses from questionnaires)
- Allow greater precision and control … but …
  are not always easy to generalize to real situations

Exercise: Interaction alternatives

Goal: Consider alternative design possibilities

Procedure
- Focus on a specific aspect of the interaction
- Create three different alternatives:
  - Predict which you think will work best
  - What are the advantages and disadvantages of each?
- We will operationalize this into an experiment

Exercise: Creating an Experiment

Goal: Choose the best interaction technique
  among several alternatives

Procedure
- Identify the key independent variables (factors)
- Identify the key dependent variables (measures)
- Operationalize the behavior and define test conditions
- Run the experiment with at least 3 subjects
- Analyze the data: are they significant?

Today

Homework: Midterm assessment (group)
  Video prototype (group)
Lectures: Evaluation: Qualitative and Quantitative
Class work: Structured Walkthrough
  Experiment design
due Jan. 15: Revise your storyboard:
  - Add three interaction points
  - Identify two potential breakdowns
  - Identify three improvements