Chapter 14
Evaluation studies: From controlled to natural settings
The aims:

• Explain how to do usability testing
• Outline the basics of experimental design
• Describe how to do field studies
Usability testing

• Involves recording performance of typical users doing typical tasks.
• Controlled settings.
• Users are observed and timed.
• Data is recorded on video & key presses are logged.
• The data is used to calculate performance times, and to identify & explain errors.
• User satisfaction is evaluated using questionnaires & interviews.
• Field observations may be used to provide contextual understanding.
Experiments & usability testing

- Experiments test hypotheses to discover new knowledge by investigating the relationship between two or more things – i.e., variables.
- Usability testing is applied experimentation.
- Developers check that the system is usable by the intended user population for their tasks.
- Experiments may also be done in usability testing.
Usability testing & research

**Usability testing**
- Improve products
- Few participants
- Results inform design
- Usually not completely replicable
- Conditions controlled as much as possible
- Procedure planned
- Results reported to developers

**Experiments for research**
- Discover knowledge
- Many participants
- Results validated statistically
- Must be replicable
- Strongly controlled conditions
- Experimental design
- Scientific report to scientific community
Usability testing

- Goals & questions focus on how well users perform tasks with the product.
- Comparison of products or prototypes common.
- Focus is on time to complete task & number & type of errors.
- Data collected by video & interaction logging.
- Testing is central.
- User satisfaction questionnaires & interviews provide data about users’ opinions.
Usability lab with observers watching a user & assistant
Portable equipment for use in the field
A selected group of panelists are invited to participate.

They are asked to evaluate the web from their natural context, using Internet Explorer.

A robot (UZ Bar) guides the users and monitors their behavior.

Remote Usability Testing

The data is analysed and a final report is prepared.

The UZ Platform gathers and saves the data in real-time.

The users are asked to complete certain tasks and answer questions.
Mobile head-mounted eye tracker

Picture courtesy of SensoMotoric Instruments (SMI), copyright 2010
Testing conditions

• Usability lab or other controlled space.
• Emphasis on:
  – selecting representative users;
  – developing representative tasks.
• 5-10 users typically selected.
• Tasks usually last no more than 30 minutes.
• The test conditions should be the same for every participant.
• Informed consent form explains procedures and deals with ethical issues.
Some type of data

- Time to complete a task.
- Time to complete a task after a specified time away from the product.
- Number and type of errors per task.
- Number of errors per unit of time.
- Number of navigations to online help or manuals.
- Number of users making a particular error.
- Number of users completing task successfully.
Usability engineering orientation

- Aim is improvement with each version.
- Current level of performance.
- Minimum acceptable level of performance.
- Target level of performance.
How many participants is enough for user testing?

• The number is a practical issue.
• Depends on:
  – schedule for testing;
  – availability of participants;
  – cost of running tests.
• Typically 5-10 participants.
• Some experts argue that testing should continue until no new insights are gained.
Name 3 features for each that can be tested by usability testing.

iPhone 4

iPad
Experiments

• Predict the relationship between two or more variables.
• Independent variable is manipulated by the researcher.
• Dependent variable depends on the independent variable.
• Typical experimental designs have one or two independent variable.
• Validated statistically & replicable.
Experimental designs

- **Different participants** - a single group of participants is allocated randomly to the experimental conditions.

- **Same participants** - all participants appear in both conditions.

- **Matched participants** - participants are matched in pairs, e.g., based on expertise, gender, etc.
Different, same, matched participant design

<table>
<thead>
<tr>
<th>Design</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different</td>
<td>No order effects</td>
<td>Many subjects &amp; individual differences a problem</td>
</tr>
<tr>
<td>Same</td>
<td>Few individuals, no individual differences</td>
<td>Counter-balancing needed because of ordering effects</td>
</tr>
<tr>
<td>Matched</td>
<td>Same as different participants but individual differences reduced</td>
<td>Cannot be sure of perfect matching on all differences</td>
</tr>
</tbody>
</table>
Field studies

- Field studies are done in natural settings.
- “in the wild” is a term for prototypes being used freely in natural settings.
- Aim to understand what users do naturally and how technology impacts them.
- Field studies are used in product design to:
  - identify opportunities for new technology;
  - determine design requirements;
  - decide how best to introduce new technology;
  - evaluate technology in use.
Data collection & analysis

- **Observation & interviews**
  - Notes, pictures, recordings
  - Video
  - Logging
- **Analyzes**
  - Categorized
  - Categories can be provided by theory
    - Grounded theory
    - Activity theory
Data presentation

- The aim is to show how the products are being appropriated and integrated into their surroundings.
- Typical presentation forms include: vignettes, excerpts, critical incidents, patterns, and narratives.
UbiFit Garden: An in the wild study
Key points

- Usability testing is done in controlled conditions.
- Usability testing is an adapted form of experimentation.
- Experiments aim to test hypotheses by manipulating certain variables while keeping others constant.
- The experimenter controls the independent variable(s) but not the dependent variable(s).
- There are three types of experimental design: different-participants, same-participants, & matched participants.
- Field studies are done in natural environments.
- “In the wild” is a recent term for studies in which a prototype is freely used in a natural setting.
- Typically observation and interviews are used to collect field studies data.
- Data is usually presented as anecdotes, excerpts, critical incidents, patterns and narratives.