Evaluation Techniques

Interaction Design Studio (DECO1200)
Outline

Observing Users
How to gather data about user observations

Interviewing Users
How to ask users questions in an interview

Designing Questionnaires
How to design different types of questions

Expert Evaluations
How to conduct expert evaluations
Observing Users

Two common lab-based techniques:

**Think-Aloud**
Users are required to say out loud everything that they are thinking and trying to do, so that their thought processes are externalised.

**Retrospective**
Users are recorded on video performing some task and then asked to comment on what they were thinking or try to do when shown the video after the session.
Data Collection

Note-Taking + Photography
Pros: Least technical methods to collect data/images
Cons: Can be tiring, limited by speed of note-taker

Audio-Recording + Photography
Pros: Less intrusive than video,
Cons: Transcribing audio to text

Video
Pros: Records visual and audio data in one
Cons: Intrusive, limited use in noisy/cluttered environs
Indirect Observation

Diaries
Users are asked to keep a diary of what they did, where they did it, and what they thought of their interactions with a product. Useful when users are scattered and not reachable in person, e.g. on the Internet.

Interaction Logging
User interactions with a product (usually software) are monitored through the use of a piece of software that captures information about key presses etc. Logged data is usually synchronised with audio/video recordings.
Interviewing Users

Guidelines for developing interviews:

Avoid long, rambling questions.
  e.g. “What was the first record that you ever owned that wasn’t give to you as a present by a family member or friend?”

Split compound questions into multiple simple questions.
  e.g. “How do you like this cell phone compared to other cell phones that you have owned in the past?”

Avoid using jargon and other uncommon language.
  e.g. “Do you use AJAX-based web services?”

Avoid leading questions.
  e.g. “Why do you like Kaz?”

Maintain a neutral position.
  e.g. “How much does Audioslave suck?”
Planning an Interview

Introduction
Interviewer should introduce himself/herself and explain why they are doing the interview, explains any ethical issues involved, asks permission to record the interview.

Warmup Questions
A set of easy, non-threatening questions are asked to help put the interviewee at ease while serving a useful purpose, e.g. gather demographic information.

Main Questions
Questions are presented in a logical sequence with the more difficult questions at the end of the session.
Planning an Interview (cont.)

Cooling-Off Questions
Finish the interview with some easy questions with the aim of diffusing any tensions that may have arisen.

Closing
Interviewer thanks the interviewee and signals that the interview is over.
Designing Questionnaires

Guidelines for developing questionnaires:

- Make questions clear and specific.
- Ask closed questions and offer a range of answers.
- Think carefully about the order of the questions.
- Avoid complex multi-part questions.
- Try to make any scales/rankings intuitive.
- Use consistent language throughout.
- Avoid jargon and uncommon language.
- Provide clear instructions for respondents.
- Avoid creating overly long questionnaires.
Sometimes the answers to a question have a limited number of alternatives.

  e.g. Questions about gender have only two answers “male” and “female”.

In such cases, the possible answers can be presented as check boxes / responses.
  Typically, a respondent will tick/cross a check box or circle a possible response.
Ranges

Some possible answers to questions are best presented as a set of ranges.

  e.g. It is common to ask people for their age, but some people don’t like to give a precise number of years. In such cases it is often better to use a set of age ranges.

When presenting ranges, ensure that it is obvious which range someone belongs to.

  A common problem is caused by providing a set of ranges like 15-20, 20-25, 25-30; but which age range does someone who is 20 belong to? It is better to provide ranges that do not overlap, e.g. 15-19, 20-24, 25-29.
Likert Scales

Likert scales are typically used for measuring opinions, attitudes and beliefs.

Likert scales are often used to evaluate user satisfaction with products.

Likert scales can be presented with numbers or with words depending on the question.

1) The use of colour is excellent: (where 1 represents strongly agree and 5 represents strongly disagree)

2) The use of colour is excellent:

strongly agree agree neutral disagree strongly disagree
Semantic Differential Scales

Semantic differential scales are used to get day about a range of bipolar attitudes. Each pair of bipolar attitudes is represented in words. The participant is asked to place a mark along a scale to indicate their agreement with the attitudes.

*Instructions:* for each pair of adjectives, place a cross at the point between then that reflects the extent to which you believe the adjectives describe the website. You should place *only one cross* between the marks on each line.

<table>
<thead>
<tr>
<th>Attractive</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Ugly</th>
<th>Confusing</th>
<th>Colourful</th>
<th>Boring</th>
<th>Pleasing</th>
<th>Unhelpful</th>
<th>Well Designed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Clear</td>
<td>Confusing</td>
<td>Colourful</td>
<td>Boring</td>
<td>Pleasing</td>
<td>Unhelpful</td>
<td>Well Designed</td>
</tr>
<tr>
<td>Dull</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dull</td>
<td>Confusing</td>
<td>Colourful</td>
<td>Boring</td>
<td>Pleasing</td>
<td>Unhelpful</td>
<td>Well Designed</td>
</tr>
<tr>
<td>Exciting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Exciting</td>
<td>Confusing</td>
<td>Colourful</td>
<td>Boring</td>
<td>Pleasing</td>
<td>Unhelpful</td>
<td>Well Designed</td>
</tr>
<tr>
<td>Annoying</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Annoying</td>
<td>Confusing</td>
<td>Colourful</td>
<td>Boring</td>
<td>Pleasing</td>
<td>Unhelpful</td>
<td>Well Designed</td>
</tr>
<tr>
<td>Helpful</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Helpful</td>
<td>Confusing</td>
<td>Colourful</td>
<td>Boring</td>
<td>Pleasing</td>
<td>Unhelpful</td>
<td>Well Designed</td>
</tr>
<tr>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Poor</td>
<td>Confusing</td>
<td>Colourful</td>
<td>Boring</td>
<td>Pleasing</td>
<td>Unhelpful</td>
<td>Well Designed</td>
</tr>
</tbody>
</table>
Administering Questionnaires

Common issues when using questionnaires include in interaction design:

- Getting a representative sample of users
  e.g. Not just the people that are easiest to ask
- Getting a significantly high response rate
  Large surveys often have 40% or fewer responses
Online Questionnaires

Online questionnaires are effective (& cheap) for reaching a large number of people.

Advantages of online questionnaires include:
- Responses are received quickly
- Costs can be very low or none
- Data can be entered into a database easily
- Data can be analysed as the results come in
- Questionnaires can be redesigned easily

E-mail vs Web-based Questionnaires

E-mail questionnaires can be easily targeted at specific people but are often limited to text only.

Web-based questionnaires are much more flexible but it can be to get a representative sample of users.
Expert Evaluations

Heuristic Evaluations
An informal evaluation technique where experts, guided by a set of usability principles (called heuristics) evaluate whether a system conforms to the principles.

Walkthroughs
Walkthroughs involve an expert performing some task with the system and noting potential usability problems.
Heuristic Evaluation

General heuristics for evaluating usability:
- Make the status of the system visible
- Match the system with the real world
- Provide users with control and freedom
- Similar actions should be handled consistently
- Help users recognise, diagnose and recover from errors
- Prevent users from making errors wherever possible
- Allow users to rely on recognition rather than recall
- Provide multiple ways to achieve tasks
- Remove unnecessary and irrelevant information
- Provide useful help and documentation
HOMERUN

Specific heuristics for evaluating commercial web sites:
- High-quality content
- Often updated
- Minimal download time
- Ease of use
- Relevant to user’s need
- Unique to the online medium
- Net-centric corporate culture
Heuristics for Web Sites

Navigation
- Avoid orphan pages
  i.e. pages not connected to the rest of the site
- Avoid long pages
  Do not force visitors to scroll unnecessarily
- Provide navigation support
  e.g. a site map or breadcrumbs
- Avoid narrow, deep hierarchical menu structures
  Studies show users prefer broad shallow menus
- Provide a consistent look and feel
  e.g. consistent navigation and hyperlink elements

Access
- Avoid complex URLs that can't be remembered
- Avoid unnecessarily long download times
Walkthroughs

Cognitive Walkthroughs
Cognitive walkthroughs involve an expert simulating a user’s problem-solving process at each step and checking that the user’s goals and memory of previous actions can be assumed to lead to the next correct action.

Pluralistic Walkthroughs
Pluralistic walkthroughs are a type of walkthrough where users, developers and usability experts work together to step through a task, discussing potential usability issues at each step of the process.
Summary

Observing Users
    Pros and cons of note-taking, audio, and video records

Interviewing Users
    Planning and conducting interviews

Designing Questionnaires
    Types of questions, check boxes, ranges and scales

Expert Evaluations
    Heuristic evaluations and walkthroughs