Lecture 6
Other Evaluation Techniques

Overview

- Other evaluation methods
  - Expert reviews
  - Field studies
- Developing scenarios
- Selecting an evaluation method

Scenarios

“A Scenario is a story” (Dumas and Redish).

- Types of scenarios:
  - Brief scenario: brief story, gives just facts of a real situation for a user, no detail on how the user goes about the task.
  - Vignettes: brief narrative may be with pictures provides broad a high-level view of user environments and current way of doing things.
  - Elaborated scenarios: provide more details.
  - Complete task scenarios: take story from beginning to end.

Scenario writing

- Can be quite complex or quite simple. Depends on the purpose of the exercise. S&P describe situations where numerous scenarios were written depending on the audience.
- Some scenarios, they argue are quite complex and involve videos to help set the scene.
- Scenarios do not need to be this complex. Often simple scenarios will suffice providing you are careful and understand the audience.

Developing scenarios and planning evaluations

- Must be mindful of the 'real' user environment in setting tasks and planning evaluations.
- Often appropriate to conduct evaluations in the environment in which the system will be used.
  Example airline check in system, environments with interruptions.
- S&P suggest a useful approach is to write scenario and then acted out as a form of theatre. Useful particularly for emergency situations.

Expert (usability experts) reviews

- Getting feedback from users, customers, colleagues.
- Can occur early or late in the design process. Take from half a day to a week.
- Outcome can be a formal report with identified problems and recommendations.
- Could result in further discussion with design teams.
- May occur at different times during development.
Expert reviews

• Long established technique. Guided by heuristics.
• Experts step through tasks role play typical users and try to identify problems.
• Expert reviewers should be placed in a situation that is similar to the real users.

“The danger with expert reviews is that the experts may not have an adequate understanding of the task domain or user communities. Experts come in many flavours, and conflicting advice can confuse the situation.” (S&P 144)
• Important therefore to choose:
  – Knowledgeable experts
  – Experts who are familiar with the project
  – Those who have longer term association with the organisation i.e. know it well.

Expert review techniques

• **Heuristic evaluation**: uses expert reviewers to critique an interface. Does it conform with the standard interface design guidelines
• **Guidelines review**: interface is checked to see that it meets the original directions/guidelines set out at the beginning of the development process

Expert review techniques - Cognitive walk-through

• Experts use the system as if they were real users and try and carry out tasks the users would perform with the system. Would begin with tasks that are performed frequently. Goal is to evaluate strengths and weaknesses of the interface.
• Walk-throughs consist of answering a set of questions about each of the decisions users are making as they use the interface.

• **Consistency inspection**: check consistency across the interface for things such as colour, language, layout workflow etc.
• **Formal usability inspection**: experts discuss interface merits and weaknesses with the designers, designers answer questions. (see S&P pg 142–144 for more detail)

• **Heuristic evaluation**
• **Guidelines review**
• **Consistency inspection**
• **Cognitive walk-through**
• **Formal usability inspection**
• Usability of the website for the target audience
• Checking that position of icons and workflow through the system are the same.
• Checking menus work and are logical.
• Checking early usability.
• Ensure original goals and design decisions were implemented
Field studies

- Aim is to observe users in their natural setting. Increases understanding of what users do and the impact of technology.
- Opportunity to collect a variety of data such as notes, pictures, audio, video, artefacts.
- Most important part of being there is to observe, ask questions and record what users are doing and saying.
- Can involve interviews.

Other considerations

- Experience of the development team in implementing different evaluation techniques.
- Type of system and experience / access of users.
- Time and $$ available
- Level of involvement of users throughout the development process.

Fitzpatrick and Dix response

Fitzpatrick and Dix (1999) argue that there are problems with these simplified matrices:
1. Uses only observational methods for real users with real computers.
2. User reports are only available when real users use representational computers.
3. Reports from specialist only obtained from representational users.

Alternative matrix

- This matrix proposes that the strategy chosen reflect the development life cycle timing.
- Strategy depends on resources available at the time for example user availability.

Evaluation methods suggested matrix

- This matrix proposes that the strategy chosen reflect the development life cycle timing.
- Strategy depends on resources available at the time for example user availability.
Dumas and Redish (1994) conclusions on evaluation methods

Conclude from the various studies that:
– Usability testing uncovers more usability problems, finds more global, unique and local problems than other evaluation methods.
– Usability testing takes more time but is more cost-effective overall.
– Heuristic evaluation done by usability specialists is better at finding usability problems than walk-throughs.
– Heuristic evaluation more valuable when used with several usability experts working independently.
– Cognitive walk-throughs least effective.
– Software engineers very poor at uncovering usability problems.

Other considerations/ issues

• Often conducted with people who are not real users.
• Fails to take into account user goals and motives
• Scenarios based on system functions not actual user needs and activities.
• Often language used in the test is not language typical users are familiar with.
• Lack of enthusiasm by the users
• Does not take into account users long-term use of the system.
• Focuses on the system not the users.

Test/evaluation Facilities

• Use of multi-purpose rooms (labs, portable computers, empty room)
• Full usability labs (see next slide)
• Opportunistic usability testing (do mini usability tests when/where ever);
• Quick and dirty (convenience rather than ‘ideal’ users).

Time and money

• Usability tests cost money and take time.
• Can take up to half a day per user.
• To make most of limited number of people must
  – Design carefully characteristics most important so defined subgroups will be useful
  – Collect other relevant information to help account for other differences that show in results
  – Choose people representative of full subgroup with full range of qualifications.

Some final thoughts

• Common principles often mentioned to contribute to improved system design:
  – Create multidisciplinary teams and involve them in all stages.
  – Understand the product life cycle.
  – Incorporate lessons learned from predecessor systems into new designs.
  – Include all team members in the planning phase.
  – Understand stakeholder needs and operations early.
  – Involve users in the design process.

• Improved product design continuously throughout its life cycle.
• Obtain management commitment and support.
Further reading and references

- Nielsen, J ‘Usability heuristics’ (1993)
- Apple Computer ‘Human interface guidelines’ (1987)